WORKING
PAPER

Pay-for-Performance in California's Workers' Compensation Medical Treatment System

An Assessment of Options, Challenges and Potential Benefits

BARBARA O. WYNN AND MELONY E. SORBERO

WR-512-ICJ

August 2007

Prepared for the Commission for Health and Safety and Workers' Compensation

INSTITUTE FOR CIVIL JUSTICE and
RAND Health

PREFACE

This paper focuses on the potential for creating financial incentives to encourage and reward the delivery of high-quality, efficient care to California's injured workers. Recently, financial incentives or "pay-for-performance" (P4P) mechanisms have rapidly gained favor in other health-care sectors but have been rarely used in workers' compensation (WC). Drawing on the models and lessons learned in group health programs, the paper assesses the options, challenges, and potential benefits of adopting P4P incentives for physician services in the California's WC program.

The RAND Institute for Civil Justice and RAND Health, divisions of the RAND Corporation, prepared the paper. It is one output of a broader study to examine selected issues in medical care provided under the California WC system. The paper was prepared for the Commission on Health and Safety and Workers' Compensation, California Department of Industrial Relations.

THE RAND INSTITUTE FOR CIVIL JUSTICE AND RAND HEALTH

The mission of RAND Institute for Civil Justice (ICJ) is to improve private and public decisionmaking on civil legal issues by supplying policymakers and the public with the results of objective, empirically based, analytic research. ICJ facilitates change in the civil justice system by analyzing trends and outcomes, identifying and evaluating policy options, and bringing together representatives of different interests to debate alternative solutions to policy problems. ICJ builds on a long tradition of RAND research characterized by an interdisciplinary, empirical approach to public policy issues and rigorous standards of quality, objectivity, and independence.

ICJ research is supported by pooled grants from corporations, trade and professional associations, and individuals; by government grants and contracts; and by private foundations. ICJ disseminates its work widely to the legal, business, and research communities and to the general public. In accordance with RAND policy, all ICJ research products are

subject to peer review before publication. ICJ publications do not necessarily reflect the opinions or policies of the research sponsors or of the ICJ Board of Overseers.

Information about ICJ is available online (http://www.rand.org/icj/). Inquiries about civil justice research projects should be sent to the following address:

Robert T. Reville, Director
RAND Institute for Civil Justice
1776 Main Street
P.O. Box 2138
Santa Monica, CA 90407-2138
310-393-0411 x6786
Fax: 310-451-6979
Robert_Reville@rand.org

RAND Health is one of the largest private health research groups in the world. More than 220 projects are currently underway addressing a wide range of health care policy issues. The research staff of more than 180 experts includes physicians, economists, psychologists, mathematicians, organizational analysts, political scientists, psychometricians, medical sociologists, policy analysts, and statisticians. Many staff have national reputations. As part of RAND, RAND Health draws on the expertise of the entire RAND staff. The program's capabilities are further broadened by long-standing collaborative relationships with other research organizations, including the University of California, Los Angeles (UCLA) and the local region of the Department of Veterans Affairs. RAND Health staff includes many physicians with joint appointments at the UCLA Medical Center and/or the Department of Veterans Affairs.

Information about RAND Health is available online (http://www.rand.org/health/). Inquiries about health research projects should be sent to the following address:

Robert H. Brook, Director
RAND Health
1776 Main Street
P.O. Box 2138
Santa Monica, CA 90407-2138
310-393-0411 x7368
Robert_Brook@rand.org

CONTENTS

FIGURES

TABLES

SUMMARY

PURPOSE

Over the past few years, nonoccupational group health-insurance programs and health plans have implemented initiatives to improve the quality and efficiency of care through incentive programs, typically called "pay for performance," or P4P. In addition, Medicare program administrators are evaluating how P4P incentives might be incorporated into Medicare payment systems. This paper assesses the options, challenges, and potential benefits of adopting P4P incentives for physician services in California's WC program.

WHAT IS PAY FOR PERFORMANCE?

Existing P4P programs reward health-care providers based on their performance on a set of specified measures related to one or more of the following:

- quality
- efficiency
- administrative processes (e.g., timely submission of reports)
- IT adoption (e.g., electronic billing)
- patient satisfaction.

A nationally prominent example is the program sponsored by Integrated Healthcare Association (IHA), a collaborative, statewide initiative that a leadership group of California employers, health plans, and physician organizations developed to stimulate improvement in patient satisfaction and clinical quality. The program involves seven health plans and 225 physician groups representing 35,000 California physicians who are eligible to receive bonus payments for individually attaining evidence-based performance goals in three areas: clinical measures, patient experiences, and investment in IT. In addition the California Office of the Patient Advocate (OPA) reports physician-group scores on patient-experience measures (see OPA, 2006). IHA recently announced that it is adding efficiency measures for episode-based and population-based group efficiency, efficiency by clinical area, and generic-drug prescribing (IHA, 2007).

WHY CONSIDER P4P FOR WC MEDICAL CARE?

Recent changes in the WC medical treatment system have created an environment that might both support and benefit from P4P initiatives. Recent statutory changes establish evidence-based treatment guidelines as presumptively correct care, require that injured workers of employers with medical provider networks use network providers throughout the course of their treatment, and authorize a new physician fee schedule.

- The medical provider networks create greater opportunity to measure performance and use both financial and nonfinancial incentives to reward providers who deliver high-quality care efficiently.
- Physicians who practice in accordance with the medical treatment guidelines could benefit from more targeted utilization review (UR) and reduced administrative burden.
- The new fee schedule creates an opportunity to align financial incentives with improved processes of care and to pay explicitly for services such as work-related disability-management activities.

In addition, the California Division of Workers' Compensation (DWC) has implemented a new data-collection system for medical services that has the potential to consolidate physician data across payers and to provide comparative data on both physicians and medical networks.

Currently, there is a high level of contention and distrust among the various stakeholders in the WC system. If properly structured to obtain the commitment and involvement of all interested parties, a P4P mechanism could be an effective vehicle for identifying common goals and improving communication and understanding among stakeholders.

WHAT ARE THE LESSONS LEARNED FROM GROUP HEALTH PROGRAMS?

Relatively little is known about the impact of the group health P4P programs or the best way to design these programs, either from the literature or conversations with individual program administrators. Substantial experimentation and refinement are occurring as programs

mature and evolve, and few have been rigorously evaluated. Nevertheless, group health P4P programs have provided valuable lessons for program implementation that may be transferable to WC. Key lessons include the following:

- Physician buy-in to the P4P program is essential. Payers must interact with and engage providers from the beginning of program development by, for example, involving physicians in the measure selection and development processes.
- All aspects of the P4P program, and most especially the measures, must be pilot-tested to work out the "bugs" and allow physicians to become familiar with any additional reporting requirements before the data are used for payment.
- P4P programs take more time and resources than initially anticipated to support the data-collection and evaluation processes. One solution is to start small and gradually build the infrastructure to support the program.

WHAT ARE THE KEY DESIGN COMPONENTS OF A P4P PROGRAM?

The most important questions that must be answered during the design and planning stages of a P4P program are the following:

- Priority goals and objectives
  - What should be the goals and objectives for a WC P4P program?
  - Should a P4P program focus on rewarding top performance or encouraging quality improvement?
  - Should a P4P program be mandatory or voluntary?
- Measures
  - What are appropriate measures for a P4P program?
  - How should the initial performance areas subject to P4P be identified?
  - What is the level of accountability for measurement and reward?
  - How is care attributed to providers?

- Data infrastructure
  - Should data be collected individually by each payer or pooled across payers?
  - What organization(s) should perform the data-collection and performance-assessment functions?

- Reward structure
  - What should be the structure of the reward?
  - What are the criteria for receiving the financial reward?
  - How is the reward financed? What data and data infrastructure could support a WC P4P program?
  - What nonfinancial rewards could be included in a P4P program?

While an array of options might be considered in answering each of these questions, the design for a P4P program must take into account a number of obstacles. These include the complexity of the current system; the lack of clinical performance measures for common WC conditions; the absence of an established, central data-collection system and ongoing performance-monitoring system to facilitate evaluation; and the involvement of multiple payers and many physicians who treat only a few injured workers each year. For this reason, a P4P program should start with relatively straightforward measurement and reward structures, but with flexibility for enhancements over time as the measurement and data-collection systems mature.

WHAT ARE POTENTIAL MODELS FOR WC?

For illustrative purposes, Table S.1 summarizes the key features to two potential mandatory systems and one collective, voluntary system among payers.

- Model 1 is a mandatory P4P program that would require minimal changes to the existing system. We describe the system as mandatory because it would be implemented by establishing fees through California's Official Medical Fee Schedule (OMFS) (see DWC, 2007a) for specific activities such as timely filing of reports or documenting specific disability-management activities. Payers would be required to pay these fees unless

they negotiated an alternative arrangement with the affected physicians.

- Model 2 is a mandatory P4P program that would add more administrative processes to the system but would facilitate the development of a broader set of performance measures over time. The system would be mandatory for physicians: All physicians relevant to the selected measures would have their performance measured automatically and would not be given the choice to participate. Data for performance measurement would be collected through the WC information system (WCIS) (see DWC, 2007b), and uniformly determined rewards would apply to high performers. Payers would be required to pay the resulting financial reward unless they negotiated an alternative arrangement with the affected physicians. Unlike the simpler fee-for-service (FFS) model, this structure would provide an opportunity for pooling data and providing report cards at the medical-group or network level when there are enough observations. Payers would also be able to supplement the financial reward structure, with additional financial or nonfinancial rewards such as reduced UR.
- Model 3 is a voluntary program modeled after the IHA P4P program, in which multiple payers voluntarily join together to create a collaborative. The payers involved in the collaborative would use a standardized set of core measures, with individual payers determining their reward structure and having the option to use additional measures as desired. Data for the core set of measures would be pooled to increase sample size for individual physicians.

We do not discuss a voluntary individual-payer model, because payers have the capability to negotiate P4P contractual arrangements with providers without additional policy development and the particulars of such arrangements would be specific to the payer and providers. However, individual-payer programs may be the most feasible to implement in the short run and could serve as pilot tests for broader programs.

Table S.1. Comparison of Three Potential WC P4P Models

Characteristic	**Model 1**	**Model 2**	**Model 3**
Type	Mandatory	Mandatory	Voluntary
Performance measures	Performing specific activities	Meeting relative or absolute thresholds for a set of performance measures	Core set of measures collectively determined
Data-collection infrastructure	Current billing system	WCIS	California Workers' Compensation Institute (CWCI) or independent organization
Reward structure	Explicit fees for work-related disability-management activities	Higher fees for "gold-star" physicians	Negotiated between payers and participating physicians
Source of funding	Employers or payers	Employers or payers	Negotiated between payer and participating physicians

NEXT STEPS

There are challenges to implementing a WC P4P, including the lack of clinical measures for WC conditions, multiple payers, and the many physicians who treat only a few WC patients. This paper offers three models that we believe might be able to surmount these problems, provided that the stakeholders have the commitment and trust to work through the design issues and allow the P4P program to evolve over time. Given the current WC environment and amount of change that has occurred in the medical treatment system over the past few years, the various stakeholders need to confirm their willingness to undertake a collective P4P initiative. Thus, a critical next step is to expand the discussion to include representatives of the various stakeholder constituencies to gauge the level of interest and commitment in a P4P initiative, define the program's goals and objectives, and determine whether there are any "idea champions" to promote the P4P concept. If there is sufficient interest, a structure could then be established for further collaborative work on design options and issues. If there is no interest in a collective effort, there may nevertheless be interest in pilot

projects to ascertain whether a P4P program would be a cost-effective way to improve the quality or efficiency of medical care. Further, there may be interest in workshops and joint educational activities on ways that individual employers or payers could incorporate P4P principles into their relationships with physicians.

While the current California WC environment poses challenges to advancing the P4P concept, this is an opportune time for DWC to put mechanisms in place to measure and reward physician performance. The opportunities include the following:

- Data collection has been initiated for WCIS, and the system should become operational by the end of 2007. There is an opportunity to consider how the medical data collected through WCIS can be turned into useful information to support performance evaluation. The critical decision is whether WCIS will support monitoring and evaluation at the system level only or whether it will be structured to also support measures of performance at the individual, group, or medical provider-network level.
- DWC has begun work to modify the physician fee schedule. As the goals for the new fee-schedule structure are established, consideration should be given to how to align the financial incentives inherent in the fee schedule with value-based care.

P4P alone will not be sufficient to drive value-based medical care provided to injured workers; rather, it should be considered as part of a multipronged set of strategies designed to increase the efficient delivery of high-quality care that enables rapid and sustained return to work.

ACKNOWLEDGMENTS

We are grateful for the support and guidance that we received throughout this study from our project officer, Christine Baker, executive officer of the Commission on Health and Safety and Workers' Compensation (CHSWC). We also appreciate the time and energy provided by others at CHSWC in supporting the roundtable discussion, including Lachlan Taylor, Irina Nemirovsky, Janice Yapdiangco and Chellah Yanga.

We are indebted to the stakeholders and knowledgeable observers of the California WC system who shared their insights and experiences with the medical treatment system in telephone interviews or participation in the roundtable discussion. We also benefited greatly from the comments of Teryl Nuckols Scott, Soeren Mattke and Thomas Wickizer on an earlier version of this paper. Finally, we appreciate the support given by Lisa Bernard, Susan Gates, Michelle Platt, and Laura Zakaras in bringing the paper to publication.

ABBREVIATIONS

ACOEM	American College of Occupational and Environmental Medicine
BTE	Bridges to Excellence
CAHPS	Consumer Assessment of Healthcare Providers and Systems
CAS	Consumer Assessment Survey
CHSWC	California Commission on Health and Safety and Workers' Compensation
COHE	center for occupational health and education
CPT	*Current Procedural Terminology*
CPT-4	*Current Procedural Terminology*, 4th edition
CWCI	California Workers' Compensation Institute
DWC	California Division of Workers' Compensation
E/M	evaluation and management
EDI	electronic data interchange
FFS	fee for service
HIPAA	Health Insurance Portability and Accountability Act of 1996
ICJ	RAND Institute for Civil Justice
IHA	Integrated Healthcare Association
IOM	Institute of Medicine
NCQA	National Committee for Quality Assurance
NOC	not otherwise classified
OHS	Washington Occupational Health Services
OMFS	Official Medical Fee Schedule
OPA	California Office of the Patient Advocate
P4P	pay for performance

TRHCA	Tax Relief and Health Care Act of 2006
UCLA	University of California, Los Angeles
UR	utilization review
WC	workers' compensation
WCIRB	Workers' Compensation Insurance Rating Bureau of California
WCIS	workers' compensation information system

CHAPTER ONE. INTRODUCTION

PURPOSE

A recent RAND study evaluating the medical care provided under California's workers' compensation (WC) system made several recommendations to drive improvements to value-based medical care provided to injured workers (Wynn et al., forthcoming). By *value-based care*, we mean the efficient delivery of high-quality care that improves the health and functional status of injured workers and enables rapid and sustained return to work. The study recommendations were to

- establish an ongoing monitoring system to assess system performance: access, quality, cost, utilization, and patient satisfaction.
- develop clinical criteria to measure appropriate care.
- make evidence-based treatment guidelines on common WC conditions and modalities readily available.
- implement a new physician fee schedule and create financial incentives to improve quality and efficiency of care.

This paper focuses on the last recommendation: to create financial incentives to encourage and reward the delivery of high-quality, efficient care. Recently, financial incentives or "pay-for-performance" (P4P) mechanisms have rapidly gained favor in other health-care sectors but have been rarely used in WC. Drawing on the models and lessons learned in group health plans, this paper assesses the options, challenges, and potential benefits of adopting P4P incentives for physician services in California's WC program.[1]

PAY-FOR-PERFORMANCE MECHANISMS: AN OVERVIEW

Over the past few years, employer group-health plans have implemented initiatives to improve the quality of care through incentive

[1] We focus on existing P4P programs for physicians in group health programs, because they offer the most formula-driven examples of P4P programs and there is publicly available information about the design, implementation, and evaluation of some of these programs. The Medicare program is testing P4P models for both physician and hospital services.

programs, typically called P4P. Existing programs reward health-care providers based on their performance on a set of specified measures related to one or more of the following:

- quality
- efficiency
- administrative processes (e.g., submission of WC-specific reports)
- IT adoption (e.g., electronic billing)and
- patient satisfaction.

As discussed in greater depth in Chapter Two, these relatively new programs are evolving, and, while their value has not been rigorously evaluated, early results are promising. A nationally prominent example is the program sponsored by Integrated Healthcare Association (IHA), a collaborative, statewide initiative developed by a leadership group of California employers, health plans, and physician organizations to stimulate improvement in patient satisfaction and clinical quality. The program involves seven health plans and 225 physician groups representing 35,000 California physicians who are eligible to receive bonus payments for individually attaining evidence-based performance goals in three areas: clinical measures, patient experiences, and investment in IT. In addition, the California Office of the Patient Advocate (OPA) reports physician-group scores on patient-experience measures (see OPA, 2006). IHA recently announced that it is adding efficiency measures for episode-based and population-based group efficiency, efficiency by clinical area, and generic-drug prescribing (IHA, 2007).

Washington, which has a single-payer state fund, offers an example of a WC P4P program. In July 2006, Washington started a two-year pilot project to improve access to surgical care for injured workers (Washington State Department of Labor and Industries, undated). The program aims to improve the access of injured workers to timely surgical care, reduce administrative burden for physicians, improve communication and coordination of care, and increase the likelihood of timely patient rehabilitation and return to work. The program is open to up to 330

orthopedic, hand, and neurological surgeons who meet the eligibility requirements. Participating physicians receive incentive payments based on their performance on a set of quality and efficiency measures. As part of the program, participating providers commit to treating a minimum of 10 injured workers during the first nine months of participation in the program; providing semiannual reports on the time elapsed between referral and completion of new patient or consultation visits; completing an activity prescription for each worker at their first visit, at the time a surgical decision is made, following a surgery, and whenever the worker's status changes; performing surgery within 21 days of receiving authorization from the claim manager; endorsing the state's preferred-drug list; and participating in occupational-health continuing education. As of July 2007, 176 providers were participating in the pilot program.

Another model of particular interest is a recently launched P4P program for spinal care. The National Committee for Quality Assurance (NCQA) has established a Back Pain Recognition Program to acknowledge physicians and chiropractors who meet performance criteria on 16 measures. The program includes measures of overuse (e.g., appropriate imaging for acute back pain) as well as measures of underuse (e.g., advice against bed rest) (NCQA, 2007). Under Bridges to Excellence's (BTE's) Spine Care Link Program, providers receiving recognition through NCQA's program receive an annual bonus payment of up to $50 for each back-pain patient covered by a participating employer. High-performing providers are also recognized on HealthGrades' physician-quality rating Web site (BTE, 2007).

CALIFORNIA'S WORKERS' COMPENSATION PROGRAM: AN OVERVIEW

California's WC program provides medical-care and wage-replacement benefits to workers suffering on-the-job injuries and illnesses. Injured workers are entitled to receive all medical care reasonably required to cure or relieve the effects of their injuries with no deductibles or copayments. It is a no-fault system, paying benefits without the need to determine whether employer or employee negligence caused the injury.

This structure is intended to ensure that workers receive prompt medical attention and needed income protection, while shielding employers from liability for civil damages and costly litigation over responsibility for workplace accidents. WC insurance covers nearly 15 million workers in California. Insurance policies issued by about 100 private, for-profit insurers and one public, nonprofit insurer cover approximately 80 percent. Self-insurance covers the remaining 20 percent (DWC and Bickmore Risk Services, 2006; CHSWC, 2006).

About 500,000 claims are filed each year for WC benefits related to workplace injuries and illnesses. About two-thirds are medical-only claims, requiring only medical treatment. In the remaining claims, the worker receives additional benefits, including temporary-disability benefits, permanent-disability benefits, and supplemental job-displacement benefits (WCIRB, 2007b).

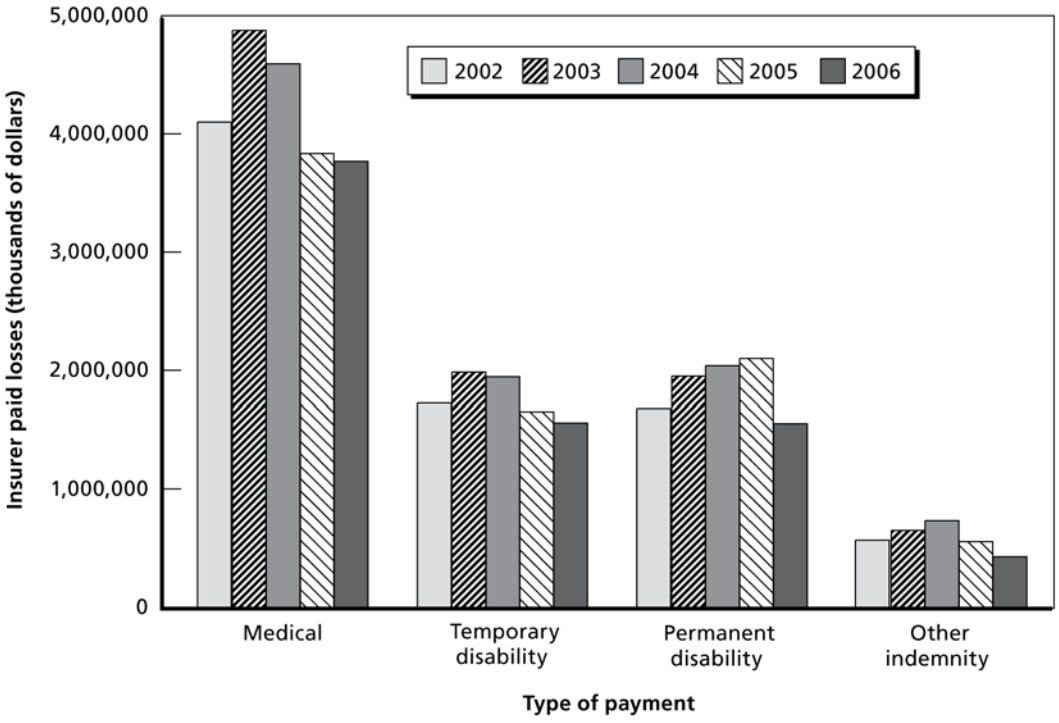
California's system for delivering WC medical care involves a primary treating physician who has responsibility for care of the injured worker. In addition to providing medical services, the primary treating physician has a central role in determining whether the worker's illness or injury is work related, in establishing the plan of treatment and making referrals for specialized care, and in assessing readiness to return to work. Many physicians treat only a few injured workers and are likely to be less familiar with work-related disability-management activities and administrative requirements than are specialists in occupational medicine. In this regard, a 2006 survey of physicians currently treating injured workers found that less than a third (31 percent) provided care to six or more WC patients per week (Kominski et al., 2007).

The California Division of Workers' Compensation (DWC) maintains an Official Medical Fee Schedule (OMFS) to set the maximum allowable amounts that may be paid for medical services to providers. OMFS does not apply if the employer or payer has contracted with a provider for a different payment amount.

California's WC program has been the center of intense debate and legislative activity over the past several years. Rising costs stimulated a series of reform efforts to control both indemnity payments

and medical treatment costs for injured workers and improve program efficiency. During the period leading up to the reforms in 2004, payments for medical care increased twice as fast as indemnity payments and represented 51 percent of paid losses in 2003 (Figure 1.1). Medical-care expenditures dropped significantly as the reform provisions were implemented.

Figure 1.1. California Workers' Compensation Insurer-Paid Losses, 2002–2006



SOURCE: WCIRB, 2003–2007a.

RAND *A7487-1.1*

For purposes of assessing options for P4P mechanisms, the important changes affecting medical treatment for California's injured workers were to

- adopt medical treatment guidelines as presumptively correct. Previously, the primary treating physician's medical decisions were presumptively correct.
- require that injured workers of employers with medical provider networks use network providers throughout the course of their

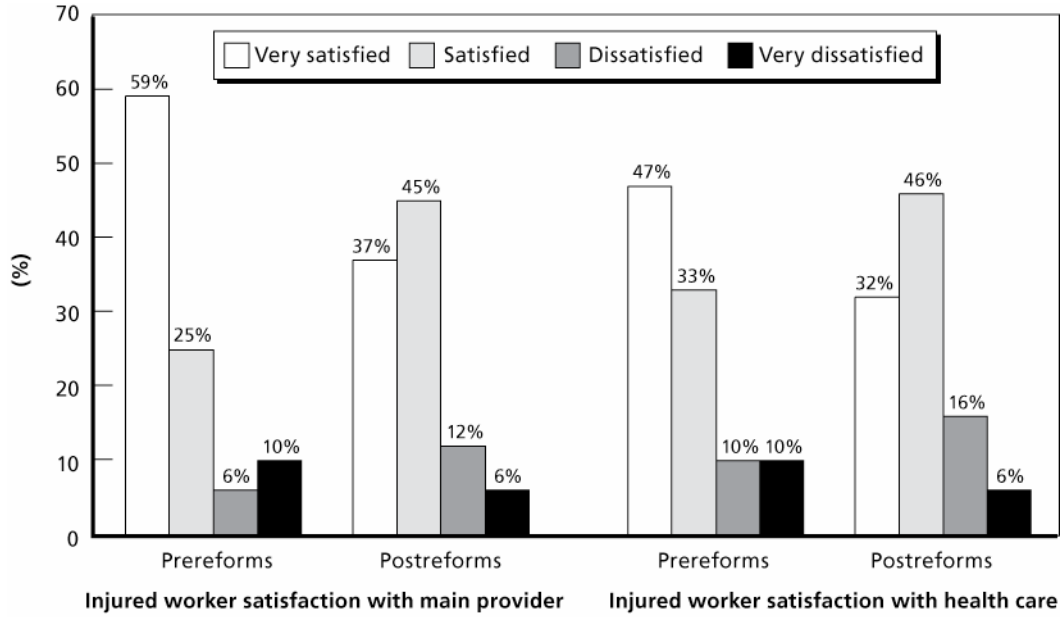
treatment. If the employer does not have a medical provider network, the prior rules that allowed the employer to control provider choice for the first 30 days and permit the injured worker to choose the primary treating physician after 30 days remain in effect.

- authorize the DWC administrative director to implement a new fee schedule for physician services. The current fee schedule is based on historical charge-based relative values that undervalue primary-care services relative to other services and do not explicitly pay for many work-related services that medical providers offer to injured workers, such as care coordination.

Relatively little is known about the quality of care provided to California's injured workers. In a four-state survey of injured workers conducted prior to the reforms, California workers did not have more severe injuries compared to workers in other states but had higher costs and lower rates of return to work, took longer to return to work, and were less satisfied with their medical care (Victor, Barth, and Liu, 2003). The reform provisions have heightened the attention given to evidence-based medicine, reduced medical utilization, and eliminated some of the poorest-performing providers from the WC system. Nevertheless, a 2006 survey of providers by the UCLA Center for Health Policy Research found that only 10 percent of respondents believed that injured-worker access to quality care has improved; the remainder believed that it had stayed the same (34 percent) or declined (56 percent) (Kominski et al., 2007).[2]

Figure 1.2. Injured Workers' Satisfaction with Main Provider and Overall Health Care

[2] Providers most affected by the reform provisions—chiropractors, acupuncturists, and orthopedic surgeons—were significantly likelier to report perceived declines in access. Nevertheless, 30 percent of physicians in internal medicine and family-practice specialties reported that quality had declined.



NOTE: The prereform data are from a survey of injured workers with a date of injury between May 1999 and December 1999, where all workers surveyed had more than seven days of lost time (Victor, Barth, and Liu, 2003). The sample of injured workers participating in the UCLA survey included workers who had a claim for injury between April 1 and June 30, 2005 (Kominski et al., 2007).

RAND *A7487-1.2*

In addition to clinical quality, patient satisfaction is an important measure of how well a medical system is performing. Patients less satisfied with their treatment are likelier to receive time-lost compensation six or 12 months after filing a WC claim than are more satisfied workers (Wickizer et al., 2004). The 2006 UCLA survey found that overall levels of satisfaction among injured workers decreased only slightly (from 84 percent to 82 percent for satisfaction with main provider and from 80 percent to 78 percent for satisfaction with health care) from a 1999 survey. Among those satisfied, fewer are very satisfied; among those who are dissatisfied, fewer are very dissatisfied (see Figure 1.2). Among injured workers dissatisfied with their health care, the three most frequently cited reasons were that they did not get the care they needed, their condition did not improve, and they did not like their provider.

WHY CONSIDER P4P FOR WORKERS' COMPENSATION MEDICAL CARE?

In the past, there has not been a strong interest in using performance measurement to improve the quality and efficiency of care provided to California's injured workers. However, the recent changes in the WC medical treatment system have created an environment that might both support and benefit from P4P initiatives.

- A common theme raised by knowledgeable individuals that we contacted in preparing this paper (see section below on methods) is that, even with the recent reforms, both the quality of clinical care and work-related outcomes have substantial room for improvement.
- Previously, employers and payers had limited ability to hold providers accountable for the quality of care and outcomes. Now that employers can establish medical networks and can control which providers care for an injured worker, they need access to performance data if they are to make wise selections for their medical provider network. The medical provider networks create greater opportunity to measure performance and use financial incentives to reward providers who deliver high-quality care efficiently.
- The authorization for a new physician fee schedule creates an opportunity to align financial incentives with improved processes of care and paying explicitly for services such as involvement in work-related disability-management activities.
- The administrative burden of complying with the new utilization-review (UR) processes has potentially affected the willingness of some physicians to treat injured workers. Performance measurement could lead to more targeted UR and reduced administrative burden for physicians with favorable performance on a set of key measures. A proactive system based on report cards may be less costly for employers and payers than comprehensive UR.
- Reliable performance measurement requires a sufficient number of patients. Because there are multiple payers and many physician practices see only a few injured workers each year, most payers do not have sufficient data to measure performance at the

physician level. While network performance is a way to address this issue, employers do not have access to comparative data across networks. DWC has implemented a new data-collection system for medical services that has the potential to consolidate physician data across payers and to provide comparative data on both physicians and medical networks.

- The various stakeholders have high levels of contention and distrust in the current WC system. If properly structured to obtain the commitment and involvement of all interested parties, a P4P mechanism could be an effective vehicle for identifying common goals and improving communication and understanding among stakeholders.

ORGANIZATION OF THIS PAPER

Technical design issues would need to be resolved in designing a P4P system for the California WC system. Chapter Two briefly summarizes the findings from our review of the literature evaluating earlier P4P initiatives involving physician services, then discusses the issues and options for key design components of a WC program, including identifying the goals and objectives for the P4P program, how to attribute care to a particular physician, how to measure performance, what type of risk adjustment is needed to avoid penalizing physicians who treat the most complex cases, and how to obtain the data needed for measurement. Chapter Three synthesizes the various options into three potential models that might be considered for a P4P initiative within the California WC system and suggests what might be the appropriate next steps in evaluating the advisability and feasibility of each model.

DATA AND METHODS

We undertook three activities in developing this paper. First, we reviewed and updated information that RAND researchers gathered for an earlier study pertaining to non–WC-related physician P4P initiatives (Sorbero, Damberg, et al., 2006) and expanded it to include available information on WC-related initiatives. Second, we conducted one-hour telephone interviews with experts involved in the California WC medical

treatment system from different perspectives: self-insured employers, payers, providers from different specialty groups, applicants' attorneys, state regulators, WC appeal judges, and researchers. Some of the interviews were with individuals, and others were with a group of individuals from the same organization. In addition, we spoke to one expert who was not involved directly in the California WC system but who was familiar with the California WC program and with efforts in other states' WC programs to improve the quality of care. We used a semistructured protocol, in which we asked interviewees about their perceptions of reform provisions' effects on medical care provided to California's injured workers, their overall assessment of whether workers have adequate access to appropriate care, and whether areas of weakness in the current system might be targeted in structuring a P4P initiative. We used the literature review and interviews to identify the key design components, options, and issues discussed in Chapter Two.

Our third activity was to convene with California Commission on Health and Safety and Workers' Compensation (CHSWC) staff a three-hour roundtable discussion with 10 California WC experts whom we identified from our interviews and through discussion with CHSWC staff. The purpose of the roundtable was to develop a better understanding of P4P issues in the California WC context and to discuss options for how a P4P program might be designed. The participants included representatives from several medical associations, payers, and the government. Seven of the 10 participants were physicians, at least four of whom provide medical care to WC patients. The two RAND researchers presented an overview of the design components of a P4P program and the findings from the literature review using a Microsoft® PowerPoint® presentation. A semistructured discussion followed on each of the major design components. For each design component, the RAND research team presented the major options identified from the literature and solicited opinions from the participants regarding how the identified options might work in the WC context and whether there were other options that might be considered. The final discussion topic was designed to learn each participant's overall assessment of how a P4P program might operate in the current California WC environment. Written notes from the roundtable

discussion were used to inform the assessment of the various options in Chapter Two and to develop the potential P4P models discussed in Chapter Three. We also consulted separately with two physicians who were invited to the roundtable but were unable to attend.

CHAPTER TWO. STRUCTURING A PAY-FOR-PERFORMANCE PROGRAM FOR WORKERS' COMPENSATION: DESIGN ISSUES AND OPTIONS

In this chapter, we describe the structure of P4P programs (Figure 2.1) and present the key design considerations that are the foundation of all P4P programs. We draw from the experience of private-sector medical-care P4P programs as well as discussions with WC experts and stakeholders to understand whether and how various options might apply to a WC P4P program. To provide a background for this discussion, we first summarize what is known from the literature on the effect that P4P programs targeted at physicians (which are predominantly based on group health programs) have on health-care quality and describe the limited identified literature specific to WC. Next, we highlight some important lessons learned that were drawn from discussions with 20 group health P4P program sponsors about their P4P programs that were conducted as part of a separate project (Sorbero, Damberg, et al., 2006) that have implications for the design and implementation of a WC P4P program.

SUMMARY OF EMPIRICAL LITERATURE ON PAY-FOR-PERFORMANCE PROGRAMS

Recent articles and reports (Dudley et al., 2004; Rosenthal and Frank, 2006; Petersen et al., 2006; Frolich et al., 2006; Sorbero, Damberg, et al., 2006) have reviewed studies examining the effect of P4P programs targeting physicians on health-care quality. These reviews are consistent in their conclusion that the limited number of existing studies produced a mixed set of results. Some of the studies reviewed found that P4P was associated with improved quality of care, most found partially positive results with performance improving on some but not all of the included quality measures, and the remainder of the studies showed no effect from P4P. Two studies too recent for inclusion in the reviews also showed partially positive results ("Generous Provider Incentives Deliver Dramatic Returns," 2006; Levin-Scherz, DeVita, and Timbie, 2006). Furthermore, these reviews found no systematic relationship between the effect of P4P and magnitude of the incentive, structure of the incentive (bonus versus fee for service [FFS]), or recipient of the incentive (individual physician versus physician

group). One review suggested that the effect of P4P may be related to the level of patient burden required to achieve improvement on the measures being examined (Frolich et al., 2007).

The empirical studies addressed mostly small-scale, financial-incentive demonstrations of limited duration. Although several of the studies reported positive results, their designs often lacked the rigor to separate the effect of the incentive from the effect of other factors occurring in the environment. The majority of these studies focused on the delivery of preventive services or ongoing care of chronic medical conditions. As a result, the literature may not be generalizable to the potential success of a P4P program in a WC system, in which the types of conditions and injuries being treated differ substantially from those measured in the available P4P studies.

Workers' Compensation P4P Literature

The reviews discussed above did not include any studies conducted in the WC system. We have identified one study evaluating a program conducted in the WC system. Wickizer et al. (2004, 2007) described the Washington Occupational Health Services (OHS) project, which was designed to improve secondary prevention to reduce long-term disability expenses and improve worker outcomes. OHS is a community-based, delivery-system intervention being implemented in two pilot sites through centers for occupational health and education (COHEs). The goals of the program are to improve the timeliness of treatment, encourage workers' return to work, and promote best practices. The clinical conditions targeted by the program are lower-back sprains, carpal-tunnel syndrome, and fractures. The program initially focused on a set of measures that are common across all of the conditions of interest, including

- timeliness of the report-of-accident submission
- two-way communication between provider and employer
- activity prescription at each patient evaluation
- assessment of impediments to returning to work
- timeliness of access to care

- adequate specification of the work relatedness of injury on the report of accident
- continuity of care.

At the time the article was written, condition-specific quality indicators were under development.

The financial reward was structured to provide reimbursement for some previously unreimbursed activities, such as telephoning the employer, and higher fees for certain activities. In addition to the financial reward, the pilot project also involved the COHE providing continuing medical education, arranging for provider mentoring by senior clinicians, and initiating care-coordination activities to avoid delays in treatment. The sources of data for the evaluation included patient- and employer-satisfaction surveys and a health-outcome survey.

The results of this project are promising. Results of the study found that, in the first year (Wickizer et al., 2004), compared to a matched comparison group, there were small improvements in the percent of cases with a report of accident submitted within two days (from 74 percent to 78 percent) and small decreases in the percent of cases in which employers were contacted to discuss return to work (from 45 percent to 40 percent). There were, however, large increases in the number of cases for which activity-prescription reports were completed (from 11 percent to 79 percent). Furthermore, providers who saw fewer WC patients exhibited less change in performance than did higher-volume providers. Regression analyses indicated the program was associated with 4,800 days and 5,800 days of reduced disability per 1,000 injured workers treated through the COHEs at the two pilot sites, respectively (Wickizer et al., 2007). Providers who were "high adopters" of submitting accident reports within two days, communicating with employers, and completing activity reports had significantly fewer disability days than providers who were "low adopters" of these practices. The COHE pilot projects were also found to be associated with cost savings, the magnitude of which varied by site; based on an analysis of one site, these savings increased over time (Wickizer et al., 2007). Due to the structure of the pilot, it is not possible to

separate the effects of the financial incentive from other aspects of the pilot.

Lessons Learned from P4P Programs

As part of an earlier study, RAND researchers interviewed 20 group health program sponsors. Despite the limited published evidence and little formal evaluation conducted internally, the researchers found that P4P program sponsors were convinced of the benefits of their programs as part of their quality-improvement initiatives. P4P programs that are operating in private-sector group health plan market are quite varied in their design. It is not clear that a single most effective design for a P4P program exists; this may depend on the environment in which the program is being implemented. P4P program sponsors were quite consistent in what they reported as challenges and as lessons learned for program development, implementation, and maintenance. Here, we highlight topics reported as most critical for a successful P4P program. See the full report (Sorbero, Damberg, et al. 2006) for the complete findings from discussions with program sponsors.

Interaction with and engagement of providers from the beginning of program development was consistently mentioned as necessary for successful program implementation. Programs accomplished this using a variety of approaches, such as involving physicians in measure-selection and measure-specification committees, the use of focus groups to obtain ongoing provider input, input from quality committees, meetings with key providers, and physician surveys. These programs' designers viewed being open to physician suggestions as critical for buy-in. Passive, high-volume methods of communication such as newsletters were viewed as inadequate for engaging physicians.

Pilot-testing all aspects of a P4P program, but most critically the measures, was another common lesson. Some programs viewed not pilot-testing as a mistake. Pilot-testing allowed a program to work out the bugs of measure definition and specification and gave physicians experience collecting and reporting the data prior to their use for payment. Two additional items that program sponsors mentioned repeatedly

were a willingness to be flexible and change and the recognition that P4P program development will involve some trial and error.

Multiple sponsors raised the issue that P4P programs take more time and more resources to manage than were initially anticipated. Resources, both IT and personnel, are needed to support data-warehousing capabilities; data aggregation; programming and analysis; data auditing; processes for appeals and data correction; provision of performance feedback; communication with, engagement of, and support of physicians; measure maintenance; and modification of data-collection processes. One approach that programs took was to start small in terms of number of physicians or number of performance measures included and gradually build the infrastructure to support the P4P program.

While many sponsors felt that their P4P programs were successful in attaining their goals, there was agreement that P4P alone is not a panacea for today's health-care problems. It was asserted that P4P needed to be implemented as part of a multipronged set of strategies designed to change provider behavior. This is consistent with alignment of payment policies with quality improvement being just one of the many actions the Institute of Medicine (IOM) recommended to improve the quality of care delivered in the United States (IOM, 2001). Other IOM recommendations, for example, addressed the need to build a health-care information infrastructure, redesign health-care processes, and improve the availability of information for patients and their families to use in making health-care decisions (transparency).

Many programs found that providers needed support to successfully participate in P4P programs and provided this in the form of patient registries, technical support, and education. At a minimum, program sponsors stated that feedback to providers needed to be actionable. As such, providing feedback in the form of rates alone does not assist physicians in improving the care delivered.

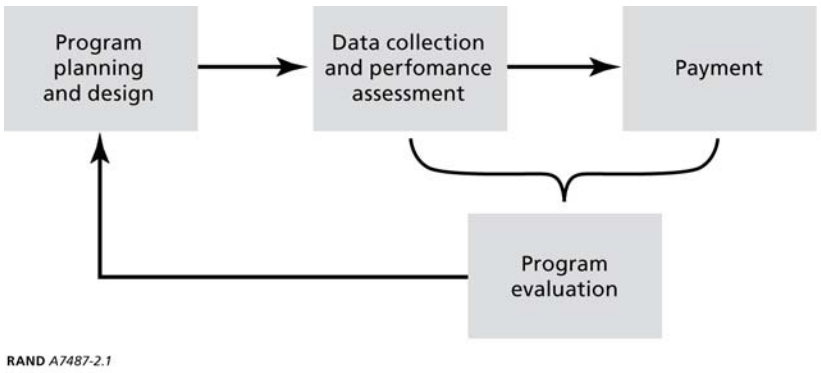
Structure and Design of P4P Programs

We conceptualize a P4P program as having four stages (Figure 2.1):

(1) program planning and design, in which the program design components are determined

(2) data collection and evaluation to assess provider performance
(3) the payment of rewards to providers
(4) program evaluation.

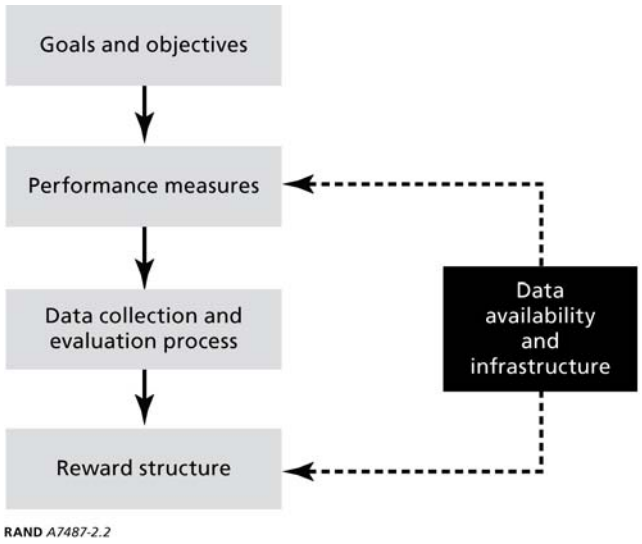
Figure 2.1. Major Stages of a P4P Program



The decisions made in the planning and design phase will structure how the program is implemented. The experiences of the program sponsor and participating providers during the operation of the program may lead to revisiting decisions made in planning and design stage and modifications in the program. Although we found from our literature review and interviews that many P4P programs are not formally evaluated, this is an important component of a well-designed program and should be performed throughout the operation of the program.

The remainder of this chapter focuses on the major components of the P4P design that need to be considered during the initial planning and design phase, including the program's goals and objectives, performance measures, the reward system, and the process that will be used for data collection and evaluation (Figure 2.2).

Figure 2.2. Major P4P Design Components



For each design component, several options could be pursued. As was highlighted in discussions with P4P program sponsors, there are no established best practices in P4P program design, and the impact of various design decisions is not well understood. Choosing among the various options typically reflects considerations of whether the approach helps achieve program objectives and what consequences may occur as a result. The choice of design options also tends to be influenced by a variety of other factors, such as the available funding mechanisms, data constraints, and stakeholder preferences. It is therefore important to clearly articulate these factors at the outset and to be mindful of them when considering various design options.

The key issues addressed in this chapter are as follows:

- Priority goals and objectives
  - What should be the goals and objectives for a WC P4P program?
  - Should a P4P program focus on rewarding top performance or encouraging quality improvement?
  - Should a P4P program be mandatory or voluntary?
- Measures
  - What are appropriate measures for a P4P program?

  - How should the initial performance areas subject to P4P be identified?
  - What is the level of accountability for measurement and reward?
  - How is care attributed to providers?
- Data infrastructure
  - Should data be collected individually by each payer or pooled across payers?
  - What organization(s) should perform the data-collection and performance-assessment functions?
- Reward structure
  - What should be the structure of the reward?
  - What are the criteria for receiving the financial reward?
  - How is the reward financed? What data and data infrastructure could support a WC P4P program?
  - What nonfinancial rewards could be included in a P4P program?

PRIORITY GOALS AND OBJECTIVES

IOM's landmark report *Crossing the Quality Chasm: A New Health System for the 21st Century* (2001) stated that health care should be safe, effective, patient centered, timely, efficient, and equitable. Consistent with these objectives, possible goals for a WC P4P program include (1) improving the quality of care delivered, (2) improving the efficiency of care delivered, (3) increasing patient satisfaction with care, and (4) improving work-related outcomes. Each of these may be further broken down.

- *Improving the quality of care* could refer to better clinical outcomes, improved processes of care, more patients getting the appropriate care for their injury, or improved care coordination among providers.
- *Improved efficiency* could reflect more cost-effective care being delivered or reduced administrative burden of the WC program.
- *Improved patient satisfaction* could reflect interactions with the primary treating provider, other providers, management of the

work-related condition, or the provider's effort to coordinate an injured worker's return to work.

- *Improved work-related outcomes* could reflect a reduction in days lost from work, a higher proportion of injured workers who experience sustained return to work in a timely fashion, or even improved productivity after returning to work.

P4P programs may have multiple goals, and the above-stated goals are not mutually exclusive. Feedback from participants at the WC P4P roundtable indicated that, while all of the above-mentioned goals are important, the primary goal of a P4P program should be to improve work-related outcomes.

Reward Top Performers or Encourage Quality Improvement by All

The decision of whether it is more important for a P4P program to reward top-performing physicians or engage all physicians and encourage quality improvement activities is partly a policy one. This decision will be one of the factors that determine how financial rewards are structured.

Programs that reward excellence set relatively strict criteria for receiving the financial incentive and rewards are distributed to a relatively narrow group of providers. Programs with this type of structure tend to reward those providers who were already performing highly prior to the start of the program (Rosenthal, Frank, et al., 2005). Poorer performers may not attempt to improve their performance if the criteria for receiving the reward seem too far out of their reach. Programs that attempt to encourage quality improvement from all providers generally are structured to distribute the rewards more broadly and set attainable criteria for receiving rewards. These two objectives do not have to be mutually exclusive, however. Programs could distribute rewards broadly but also have a larger financial incentive for those providers who excel in their performance.

Discussants at the WC P4P roundtable indicated that quality improvement should be the focus of a P4P program. This suggests that a WC program should be structured to distribute financial rewards broadly. This would engage most physicians in the program by making receipt of

the financial reward attainable; few are likely to dismiss the program for viewing the incentive as impossible to obtain. If the size of the financial reward is linked to performance, physicians have an incentive to strive for continued improvement in their performance to obtain a larger reward.

Mandatory or Voluntary Program

A mandatory WC program would require legislative or regulatory action on the part of the state to require all payers and self-insured employers to participate. Because the state would define how providers would be evaluated, select the performance measures included in the program, and select the performance targets, these would be uniform throughout the system. This uniformity would simplify participation requirements for providers and payers, facilitate the pooling of data, and potentially enable public reporting of quality information. The reward structure, however, could be defined uniformly or left to each payer to determine. An advantage of a mandatory program is that it facilitates the pooling of data (because the same performance measures are used universally) and public reporting.

A voluntary program could be undertaken by payers individually, as is the case for many physician P4P programs in other health-care sectors, or collectively. A voluntary program has the advantage of not requiring government involvement. On the other hand, individually operated programs that lack central coordination could lead to multiple reporting requirements that would increase the administrative burden for providers, possibly create conflicting signals of what it means to deliver high-quality care, and weaken the financial incentives of the individual programs. Collective action facilitates the use of pooled data. An example of a collective-action, voluntary initiative is the IHA program in which the individual health plans are using the same core set of measures and are pooling data but have constructed their own financial incentives independently.

Discussants at the WC P4P roundtable indicated that the question of whether a program should be mandatory or voluntary should be open for debate, recognizing the advantages and disadvantages of each approach.

Discussants noted that a voluntary program was likelier to be politically feasible but that it also might be possible to pilot a mandatory program in selected counties to gain experience.

WHAT TYPES OF MEASURES SHOULD BE INCLUDED?

Several types of performance measures can be chosen for reward. Current programs frequently use more than one type of measure in their programs (Sorbero, Damberg, et al., 2006). Decisions among the alternatives depend on program goals and the feasibility and costs of data collection. Measure options include clinical effectiveness (process—whether the provider is providing care according to guidelines and standards—and outcomes—end points of care, such as restoration of function), patient satisfaction, cost of care, characteristics of care systems, administrative measures, and work-related outcomes of care. Regardless of the specific types of measures included in a WC P4P program, participants at the WC P4P roundtable suggested that pilot-testing is necessary for providers to gain familiarity and comfort with the individual measures.

Option 1: Clinical Process and Outcome Measures

Clinical-process measures assess whether the patient received a recommended test or service—the patient getting the "right care at the right time." Clinical processes typically are recorded either in the medical chart or in claim data for billing purposes. Multiple national organizations (e.g., National Quality Forum, Agency for Healthcare Research and Quality, IOM) have criteria for the evaluation of clinical-process measures. The IOM criteria (2006) include (1) scientifically sound (i.e., reliability, validity, and explicitness of the evidence base for the measure), (2) feasibility (i.e., availability of data and burden of measurement on providers), (3) importance (i.e., measure focuses on a leading cause of death or disability or is associated with high resource use), (4) alignment (i.e., measure specifications are consistent across different measure sets to reduce redundancy and provider burden), and (5) comprehensiveness (i.e., measure adds to the ability of the entire measure set to address the way care is delivered).

A 2006 review found a dearth of clinical process measures suitable to occupational disorders (Nuckols, Asch, et al., 2006). The existing measures that were relevant targeted a limited number of clinical areas, such as lower-back pain, shoulder injury, and knee injury. Many of these existing clinical-process measures that are relevant to WC conditions have not been developed adequately to ensure accurate and reliable measurement; others require updating to reflect current clinical knowledge. Some progress, however, has been made since the 2006 review was conducted. The NCQA recognition program recently adopted a set of measures for certain aspects of care for lower-back pain (NCQA, 2007):

- initial visit
- physical exam
- mental-health assessment
- appropriate imaging for acute back pain
- repeat imaging studies
- medical assistance with smoking cessation
- advice for normal activities
- advice against bed rest
- recommendation for exercise
- appropriate use of epidural steroid injections
- surgical timing
- patient reassessment
- shared decisionmaking
- patient education
- postsurgical outcomes
- evaluation of patient experience.

In general, aspects of patient care that are not critical for reimbursement, such as return-to-work planning, are less likely to be documented in the medical record. This can affect the reliability of the measures used in the P4P program. For example, the initially observed improvement in performance based on medical charts may really result from improvements in documentation rather than actual changes in patient care. Beyond this NCQA work in back pain, RAND researchers are developing clinical-process measures for carpal-tunnel syndrome. These

measures are expected to be tested and validated in approximately a year.

Many of the available measures are based on medical-record review rather than administrative data, which increases the burden of data collection. Due to this data-collection burden, the use of medical records by group health P4P programs is much less common than the use of claim data or other administrative data, and data tend to be collected for only a sample of relevant patients rather than all patients eligible for a measure (Sorbero, Damberg, et al., 2006). Medical records are costly to abstract, a cost that could be borne by the entity collecting the data, be it providers, the P4P program sponsor, or shared. An alternative approach would be for physicians to collect and report additional information needed for performance measurement with the program sponsor offsetting the costs of this activity, separate from any financial rewards that physicians would receive through the P4P program. For example, the Tax Relief and Health Care Act of 2006 (TRHCA) authorized a voluntary physician-quality reporting system for the Medicare program. Eligible professionals who successfully report a designated set of quality measures on claims may earn a bonus payment, subject to a cap, of 1.5 percent of Medicare physician fee schedule payments.

Outcomes of care provide another assessment of the clinical quality of care. Clinical-outcome measures most frequently examine patient mortality but could also include rates of temporary and permanent disability, as well as duration and severity of disability. In other health-care sectors, mortality is the only routinely collected outcome, and there is no existing system for tracking patients' health over time. For this reason, and because of challenges in assigning accountability for long-term outcomes, designers of existing quality-monitoring systems in other sectors have found it much more feasible to use process-of-care measures than outcome measures.

In WC, several important disability variables are routinely tracked, which may make the inclusion of outcome measures in a WC P4P program less problematic than in group health P4P programs. Nevertheless, there are other challenges involved in the use of outcome

measures. Because factors other than quality of care can affect outcomes, such measures need to be adjusted for differences in the patient mix treated by each provider. Underlying patient characteristics that affect patient outcomes include comorbid illnesses and severity of illness; differences in patient factors shown to affect compliance, such as education, socioeconomic status, race, ethnicity, primary language spoken (DiMatteo, 2004; Bosworth et al., 2006; Kaplan et al., 2004; Sloan et al., 2004). For WC, other factors, such as employee-employer relations and employer willingness to accommodate work modifications, have been shown to affect return to work, which could impact disability measures. While there is recognition of the need to adjust for differences in patient populations, the necessary data for adjusting across different outcomes are frequently absent or require merging data from multiple sources. For example, information on education, socioeconomic status, and ethnicity is rarely collected as standard practice. For WC, this information should be part of the administrative record for an injured worker, but the administrative record and the medical data may not reside in the same database (and may be maintained by different organizations).

Furthermore, many risk-adjustment methods are complicated to implement. Adjustment is a critical consideration because of the possibility that P4P will otherwise create incentives for physicians to avoid patients who are more difficult to treat and who are expected to have a negative impact on them financially (Weissert and Musliner, 1992; Ellis and McGuire, 1996; Newhouse, 1989; Sorbero, Dick, et al., 2003; Shen, 2003). Methodological issues include the following:

- A long time frame is frequently needed to observe some outcomes. While intermediate outcomes can often be observed within three to 12 months from time of treatment, long-term outcomes may not be observed for many years, making them difficult and expensive to track. In WC, a claim may not be closed for years.
- Another significant challenge with long-term outcomes is determining how to attribute the final outcome to a specific provider's actions when there are intervening providers and events.

- Many outcomes of interest are rate events, such as deaths following care and major complications of surgery. This makes it difficult to observe these events and to detect statistically significant differences across providers.
- Most data systems that contain information on intermediate or long-term outcomes (e.g., laboratory data, death records) are not linked to existing administrative systems from which performance can be measured. Consequently, substantial investments in resources may be required to link information.

Potential areas for clinical measures suggested by participants during the WC P4P roundtable included referrals to specialists and other practitioners that were consistent with treatment guidelines and repeated surgeries on the same body part. To the extent that these areas are not yet specified as actual measures, they would need to be developed and tested prior to their use in a P4P program.

Option 2: Measures of Efficiency or Resource Utilization

The key argument for including these types of measures in P4P programs is the continued upward pressure on the costs of care in the WC system, both in terms of medical costs and indemnity costs as well as variations across providers in amounts of services used to treat patients with similar conditions. There are several options for measuring resource utilization, ranging from simple utilization measures, such as number of physical therapy visits per claim, to more complex measures of relative resource use, such as cost per episode of care. Efficiency measures that could be used in WC include total medical costs or total costs (medical and indemnity) for a provider relative to peers. One challenge that providers face in using efficiency measures is that they are not very actionable—providers frequently need assistance to understand what they need to change to improve their scores or become a more efficient provider. For example, a summary measure on total medical costs or total medical and indemnity costs does not provide a physician with information on what services might be over- or underutilized relative to other physicians treating comparable patients. More actionable information would be utilization measures reflecting use of ancillary services and referral rates for specialist services.

Participants at the WC P4P roundtable suggested that, in the WC environment, efficiency in returning an injured worker to sustained work is an important outcome that should be part of a performance-monitoring system. Participants thought that there may be trade-offs between medical and indemnity costs; for example, a more aggressive medical rehabilitation treatment plan entails higher medical expenditures but, if it also enables the worker to return to work earlier, it would be more efficient than a less aggressive course of treatment. For this reason, participants suggested that total costs would be the most appropriate measure of provider efficiency.

Option 3: Assessment of Patient and Employer Experience

Assessments of patient experience are typically based on survey information and assess provider performance associated with such issues as patient access to care, coordination of care across providers, and doctor-patient communication. These measures capture aspects of care that are important to patients and represent a different domain of performance. The national Consumer Assessment of Healthcare Providers and Systems (CAHPS) consortium has developed a series of ambulatory-care surveys for use at the physician level, which could be modified to ensure its relevance to a WC population (Farley, 2006). URAC, a national nonprofit organization that has established accreditation standards for health-care organizations, has developed a survey tool for injured workers so that WC managed-care organizations can gather information about the quality of care they deliver (American Accreditation HealthCare Commission/URAC, 2000). However, this tool has not undergone rigorous testing, and procedures for scoring the results of the survey have not been developed. Both of these steps are necessary to use the tool to construct a valid measure of patient experience for P4P and to determine appropriate methods for including the measure in a P4P program. While the inclusion of patient-experience measures is advantageous in that all physicians could be measured and patient-centered care is a key dimension of care, these measures could not be produced without new data being collected at the direct expense of the providers or insurers.

The Ohio Bureau of Workers' Compensation uses an external organization to survey both injured workers and employers to assess their experience with the managed-care organizations that help manage WC claims and coordinate medical services. Employers are asked how satisfied they were with the organization's service, medical management, efforts to provide appropriate and early return to work, and educational and other training materials, as well as the employer's ability to contact the organization. For employees, the questions include how satisfied they were with their ability to contact the organization when needed, choice of physicians, the medical treatment and medications received since the injury, and efforts to help them return to work and rehabilitative services.

Discussants at the WC P4P roundtable suggested other measures of patient experience that could be constructed using administrative data rather than requiring primary data collection. A leading suggestion was provider retention rates (e.g., whether patients change to other primary treating physicians). While there are factors that drive patients to switch physicians other than the actual experience in receiving care, such as proximity, switching physicians has been conceptualized as capturing both satisfaction with care and trust in the physician (Sorbero, Dick, et al., 2003).

Option 4: Administrative Measures

The WC system has more administrative requirements than providers are used to in the group health system. As many physicians care for only a small number of injured workers and physicians are not separately reimbursed for the completion of WC forms, learning and complying with administrative requirements of WC programs may not be a high priority for providers. While regulations are in place to encourage the timely submission of many forms (e.g., the primary treating physician is supposed to submit the first report of injury within five days of the initial examination), penalties for noncompliance are not systematically enforced. As a result, many forms required for WC patients are not submitted in a timely fashion or are incomplete when submitted. Examples of administrative measures that could be included in a WC program

include the timely submission and completeness of submission of specific forms by the primary treating physician, such as the doctor's first report of occupational injury or illness and the report that is required when the patient's condition is permanent and stationary. These administrative activities are important to facilitate communication with both the payer and employer and may reduce the time it takes for the worker to return to work. In addition, research has shown that delays between injury and first medical treatment as well as longer time from medical treatment to claim filing are predictive of long-term disability (Stover et al., 2007). As a result, adherence to administrative measure may help reduce the overall costs of an injury.

Participants in the WC P4P roundtable discussion suggested that the following measures may be appropriate for inclusion in P4P: (1) triggering of noncertification of referral (i.e., patient is referred to another provider without first obtaining certification), (2) claims that are litigated, (3) provider writing a modified duty prescription for the patient, and (4) provider performing work activity assessments. Much of this is information that the WC program already has in administrative files or is information that claim administrators request from providers.

Option 5: Work-Related Outcome Measures

Work-related outcome measures do not suffer from some of the challenges with health outcome measures. Work-related outcomes, such as the injured worker returning to work, occur with greater frequency than mortality and closer to the time of injury and delivery of care, making their measurement easier. In addition, employers or payers track them, thereby reducing the administrative burden associated with their use. One methodological challenge that work-related outcome measures share with other outcome measures is the need for risk adjustment. Another challenge is the variation that exists among types of providers and types of employers on their views on facilitating return to work through work modifications. Some employers want injured workers on the job as soon as possible, while others prefer that injured workers not return until no work restrictions are needed.

The Ohio Bureau of Workers' Compensation includes an "optimal return-to-work" measure in its report card for managed-care organizations. The measure compares actual time lost with established benchmarks for loosely managed and well-managed disabilities that take into account the injury and occupation. The score is based on the organization's progress from its loosely managed benchmark to its well-managed benchmark (Ohio Bureau of Workers' Compensation, 2007).

URAC includes among its WC performance measures time to return to work, lost-time days, and total compensation days. The participants in the WC P4P roundtable also supported these measures as attractive candidates for use in a WC P4P program. Another possible measure that is less frequently focused on by WC but is of interest to employers is productivity when the worker returns to work (Mattke et al., 2007).

Option 6: Structural Measures

Structural measures represent another potential area of measurement and reward (MedPAC, 2005). Group health programs sometimes include structural measures, such as board certification and use of electronic health records and use of computerized, physician order-entry systems in the P4P programs. For WC, completion of continuing education in topics related to care of injured workers could be another type of structural measure. Rewarding IT and other measures of system support may serve to stimulate faster improvement in quality and create a business case for investment in systems. This would further accelerate topics put forth in California Governor Arnold Schwarzenegger's health-care proposal, including universal e-prescribing by 2010 and 100-percent electronic health-data exchange within 10 years (Hill, 2007). A difficulty with using IT measures is that they could be burdensome for doctors to self-report and potentially a challenge for payers to aggregate and analyze across the numerous physicians in their network.

OTHER KEY ISSUES RELATED TO MEASURES

Other key issues related to measurement in a P4P program include prioritizing the initial focus of the program, establishing the level of accountability or focus, and determining how to attribute the

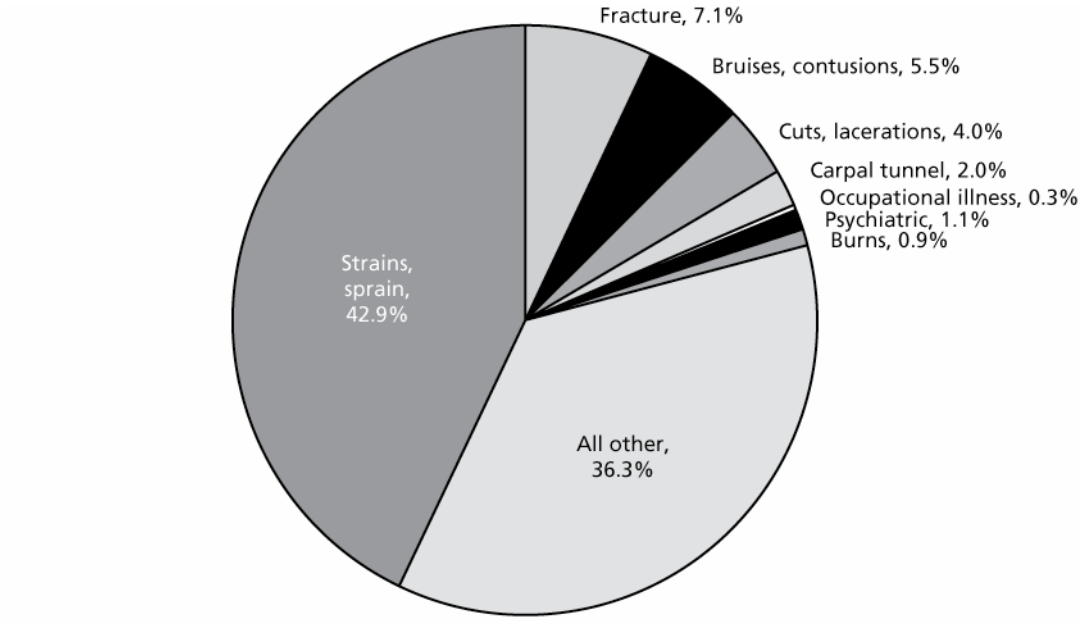
experiences of individual workers to providers. Each of these is discussed below in turn.

Initial Areas to Target

Several factors can drive the decisions about where initially to focus program-measurement and incentive efforts.

Prevalence of conditions in the WC population ensures that the areas of focus are important and increases the likelihood of having adequate numbers of patients on which to base assessments of provider performance. Figure 2.3 displays information from WCIRB on the most common conditions appearing on individual case reports involving death claims, permanent disability, or temporary disability losses of at least $5,000. Sprains and strains accounted for more than 40 percent of the diagnoses on all cases, followed by fractures (7 percent), contusions (6 percent), and lacerations (5 percent), suggesting that these may be attractive areas for measurement.

Figure 2.3. Common Conditions in Workers' Compensation Claims



SOURCE: WCIRB, based on injuries in 2004 involving death, permanent disability, or temporary disability losses of $5,000 or more (WCIRB, 2007a).

RAND *A7487-2.3*

High-cost conditions can be targeted in an effort to improve processes of care and address the overutilization of services for certain conditions. Table 2.1 displays information on the top 20 injury categories measured by total losses. Those categories that also have an average loss per claim (including outliers) that is among the top 20 injury categories are shaded. Some of these categories are nonspecific (e.g., all other cumulative injury, not otherwise classified) and would not be potential target conditions. Other conditions that have a high acute-care component, such as burns, may also be less attractive than other conditions for a WC P4P initiative focused on physician care such as sprains, strains, and carpal-tunnel syndrome.

Table 2.1. Top 20 Injury Categories for Total Incurred Losses

Nature of Injury	Claims	Incurred Losses ($ millions)	Loss per Claim
Strain	40,345	1,580	39,164
All other cumulative injury, not otherwise classified (NOC)	14,554	631	43,394
All other specific injuries, NOC	15,450	627	40,612
Fracture	7,103	373	52,531
Sprain	8,996	333	37,038
Contusion	6,801	253	37,136
Multiple physical injuries only	3,333	194	58,139
Laceration	4,825	150	30,997
Inflammation	2,975	110	36,887
Carpal-tunnel syndrome	2,242	90	40,240
Dislocation	1,084	60	55,404
Burn	699	47	67,799
Crushing	956	43	44,521
Mental stress	1,668	38	22,770
Amputation	644	37	57,474
Hernia	2,109	37	17,456
Myocardial infarction	643	25	39,392
Puncture	738	23	31,631
Multiple injuries including both physical and psychological	717	22	31,366
Rupture	371	22	58,562

SOURCE: WCIRB (2006).

Known quality deficits in an area is a frequent reason for inclusion in a P4P program. Little is known about the quality of care delivered to WC patients and how this may differ by specific clinical conditions and as such cannot be used to guide the development of a WC P4P program. While guidelines exist for many common WC conditions, relevant measures are uncommon. Before a system for improving quality will seem justifiable, measurements confirming specific quality problems will be needed.

Administrative measures that apply to all WC patients are another potential focus area. There are known delays and gaps in the submission of administrative information that can be used to facilitate an employee's return to work.

Level of Accountability

The decision about which unit of accountability or level of focus—individual physician, medical group, or network—to measure and reward is an important decision and is influenced by the program's goals as well as practical issues. For example, if a key goal of the program is to increase coordination of care for WC patients, then focusing on a medical group or network rather than on individual physicians may help promote coordination across providers. The network is also an appropriate level of evaluation if the goal is to provide employers with information about payer-created networks. If the goal is to provide payers and self-insured employers with information to enable them to build a network of high-quality doctors, measurement at the level of the individual physician or medical group might be more appropriate.

If measurement is focused at the medical group or network level, rather than the individual physician level, there is an operational challenge of mapping individual providers to the medical groups or network. An initial map would need to be created as well as a process for updating it frequently. Private-sector P4P programs report that mapping individual physicians to medical groups requires considerable resources and knowledge of the local health-care market. Mapping to the

network would be more straightforward, as the payer would maintain this information as part of business activity.

Choice of the level of focus may also be influenced by statistical considerations associated with the performance measures, primarily the problem of small numbers. WC patients frequently make up only a very small proportion of an individual provider's practice. Thus, for any specific clinical measure, many providers would have fewer patients in the denominator than the minimum number (frequently 30) needed to ensure that the ability produces a stable estimate of performance. Three options that can be used to address the problem of small numbers are

1. pooling data across multiple payers or self-insured employers (though this still may not generate adequate numbers at the level of individual physicians).
2. pooling data over time (e.g., use two years of data to assess performance rather than one). A disadvantage of this approach is that providers do not receive immediate, actionable feedback that can lead to performance improvement. The delay can also weaken the financial incentives.
3. aggregating over multiple measures to create composite scoring (e.g., across all measures relevant to sprains and strains). There are issues about composite construction that are not well understood and would require further testing, however. These include what measures to combine (i.e., are individual measures related to one another in a way that will yield a clear signal of quality?), whether to weight each individual measure included in the composite equally or differentially, and what the effects of different decisions around grouping and weighting of measures would be.

Participants in the WC P4P roundtable discussion had varying views of the appropriate level of focus for a program. Contracts are between the employer or payer and the network; however, there is no mechanism in place to pay the networks—payment currently flows directly to providers under the OMFS. This suggests that the individual physician may be the appropriate focus for P4P. However, given that many physicians have a relatively small number of WC patients, it is unlikely that measurement

at the level of the individual physician will be feasible for measures requiring the construction of rates. An alternative would be to use metrics that do not require construction of rates in the P4P program, such as structural measures or provision of the reward each time a desired activity (such as timely completion of a report) is performed.

How Is Care Attributed to Providers

All P4P programs have a defined method for attributing patients (and their care) to the provider that is the unit of analysis (i.e., physician, medical group, network). Attribution rules can be structured, in some cases, to help achieve the goals of the program, such as to create incentives to coordinate care, a sense of responsibility for the patient, and physician buy-in to the P4P program. Different attribution methods may lead to different results and have different implications for the P4P program. However, there is a lack of empirical evidence about the differential effect of specific attribution rules. In light of a lack of empirical evidence, we suggest the factors that should guide the selection of an attribution methodology are that it (1) is reasonable and is perceived as fair by physicians, (2) creates a sense of responsibility for the delivery of care, and (3) is operationally feasible to implement and manage (e.g., assignment can be made automatically based on existing data). We describe here two general options for attributing patients to physicians, one of which has three variations, and highlight some of the likely effects.

Option 1: Attribute All Care to the Primary Treating Physician

The simplest approach in the California WC system would be to attribute responsibility of patient care to the primary treating physician. The rationale for doing so is that this physician has overall responsibility for the patient over the course of treatment for the injury, acts as a gatekeeper for referrals to other physicians, and performs work-related functions. However, patients whose employer has a medical provider network may change treating physicians after the first visit, and other injured workers may switch primary treating physicians after 30 days. As a result, an injured worker may have more than one primary treating physician throughout the course of treatment. Another

challenge with this method is that the primary treating physician may not, in fact, provide the specific aspect of patient care that is subject to measurement (e.g., the primary treating physician is a chiropractor and the care being measured is related to a surgery).

Option 2: Attribute Care to Those Physicians Who Touched an Injured Worker

There are variations in how this option could be constructed:

- Based on all physicians who provided care to the patient that are of the relevant specialty for the measure
- Based on all physicians who provided care to the patient regardless of specialty
- Based on the physician who provided the care of interest to the patient (e.g., an orthopedist who set a fractured bone).

The first two approaches under option 2 differ from the method described in option 1 and the third method under option 2 in that there is likelier to be shared performance responsibility, since multiple providers may be accountable for patient care. This may maximize the probability that a patient receives the recommended care, since all relevant physicians are responsible for ensuring that a patient gets the recommended care. However, the effects of this approach are not well understood; it is not known whether it actually increases or decreases physicians' sense of responsibility for care delivery. Depending on how it is constructed, this method runs the risk of double-paying on any given performance measure—that is, paying a physician who did not provide the service but who was nonetheless assigned responsibility.

Another important consideration is that the attribution method must appear reasonable to providers. Assigning patients to physicians who believe that they are not primarily responsible for those patients' care may reduce provider acceptance of the program.

DATA INFRASTRUCTURE TO SUPPORT A PAY-FOR-PERFORMANCE INITIATIVE

P4P initiatives require a data infrastructure to support core data-related functions, including

- data collection and cleaning

- data warehousing
- data analysis and performance feedback.

Individual claim administrators (self-insured employer, third-party administrators, or WC payer) maintain the internal data infrastructure needed to administer benefit determinations and payments for injured workers. However, the data needed to operate a P4P program may not reside in a single data warehouse. For example, a self-insured employer may maintain data on a claim's administrative history and indemnity payments, but detailed information on medical expenditures may reside with a third-party administrator. WC payers are likelier to maintain a single database containing a comprehensive history of the claim, but there may be separate databases for medical administrative data (which may reside with a medical-review organization instead of the payer).

Two main WC data systems collect data from multiple payers and might be used to support a P4P initiative:

- The California Workers' Compensation Institute (CWCI) is a private, nonprofit organization of payers and self-insured employers whose members voluntarily submit data to CWCI for research. General administrative data is collected as well as transaction-level information on medical and indemnity payments over the life of the claim. Building on the CWCI data infrastructure would likely work best for a voluntary P4P program.
- DWC maintains the WC information system (WCIS). Claim administrators report certain elements of their data to WCIS via an electronic data interchange (EDI). Effective for dates of service beginning September 22, 2006, payers and employers with more than 150 claims annually are required to report detailed, transaction-level information on medical services. Because it is a new reporting requirement, the consistency, quality, and completeness of the data are uncertain. Once WCIS is fully operational, it could be used in either a voluntary or mandatory P4P program.

As discussed in the sections that follow, we have identified four basic options that might be used to support the data functions of a P4P program:

- Option 1: Individual payers conduct their own data collection and evaluation.
- Option 2A: DWC collects necessary data through WCIS and evaluates the data.
- Option 2B: DWC collects necessary data through WCIS, and an independent third-party organization evaluates the data.
- Option 3: An independent organization conducts a separate data-collection effort and evaluates the data.
- Option 4: CWCI collects and evaluates the data.

Other important data-related functions need to be incorporated into a P4P program. These include

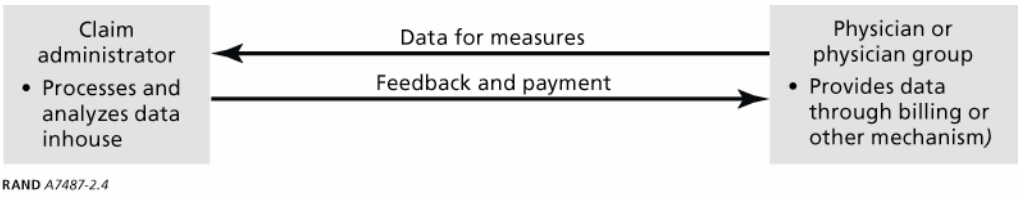
- auditing
- feedback process
- review and correction of data
- reward determinations.

How these functions are carried out will depend on how the P4P program is structured and the model adopted for the core data functions.

Option 1: Individual Payers Conduct Their Own Data Collection and Evaluation

Under this option, each claim administrator would perform its own data-related functions independent of other payers and reward performance based solely on measures derived from a provider's performance on measures related to injured workers covered by that claim administrator (Figure 2.4).

Figure 2.4. Data-Collection Process Is Internal to Individual Payers



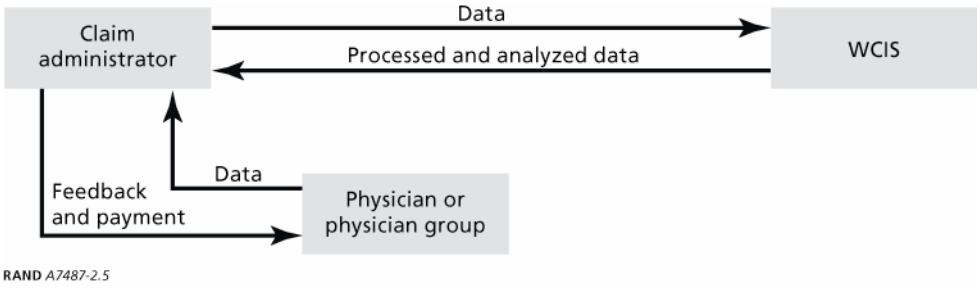
This model allows each claim administrator to decide on the level of effort with regard to P4P and reward structure. It would be most

appropriate for pilot demonstrations and voluntary P4P programs. It loses the advantages gained from data pooling and may make it difficult for smaller payers to adopt P4P incentives. However, it would not preclude voluntary adoption of a standardized set of measures across payers. Several payers have expressed an interest in using this approach for either selective contracting or P4P. In addition, some networks have also expressed an interest in collecting quality data for internal purposes.

Option 2A: DWC Collects Necessary Data Through WCIS and Evaluates It

This model (Figure 2.5) could evolve into a viable option if DWC were to develop an ongoing monitoring and evaluation system that brings together the full spectrum of data needed to assess system performance. It would be an efficient mechanism for collecting and pooling the data needed to support a P4P system, but it would also require a significant expansion of DWC staff and resources to evaluate the data and commitment from the California state government to adequately support the effort. The model assumes that a standard set of measures would be collected, pooled, and evaluated. The model could work with either a voluntary or mandatory P4P program and could give each employer the flexibility to determine its own reward structure.

Figure 2.5. Data Collection and Evaluation Processes Are Through WCIS

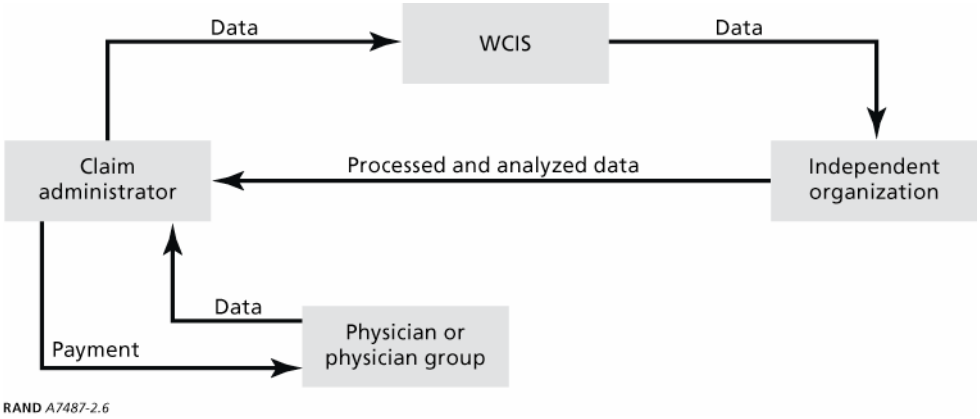


Option 2B: DWC Collects Necessary Data, but an Independent Third-Party Organization Evaluates It

Option 2B (Figure 2.6) preserves the benefits of DWC collecting the necessary data through WCIS but relies on an independent third party to evaluate the data. This model is likelier to provide sufficient staffing

and resources to ensure successful operation of P4P program, because the organization's staffing could be dedicated to this effort and have the appropriate skills and experience with performance measurement to evaluate the data. The organization's activities could be supported through appropriations or a payer assessment (in the case of a mandatory program) or through user fees (in the case of a voluntary program). As is the case with option 2A, there is flexibility to determine the reward structure uniformly or by payer.

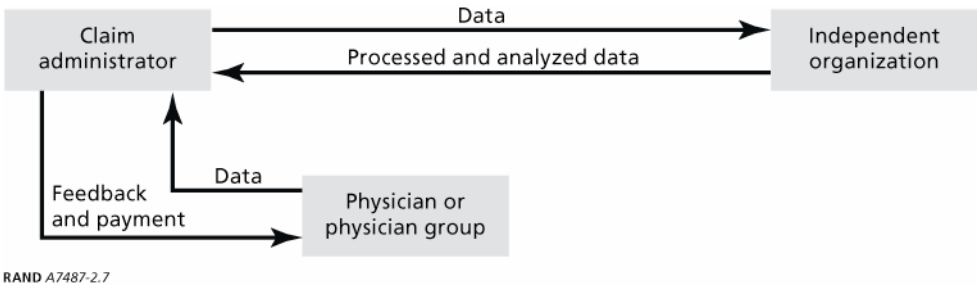
Figure 2.6. Data Collection Is Through WCIS and Evaluation Is Through an Independent Organization



Option 3: An Independent Organization Collects and Evaluates the Data

The model (Figure 2.7) contemplates establishing a separate data infrastructure to support P4P initiatives with the appropriate skill sets and relieves DWC of any administrative burden associated with the program but creates additional data submissions for claim administrators. The model would work best with a collective voluntary program.

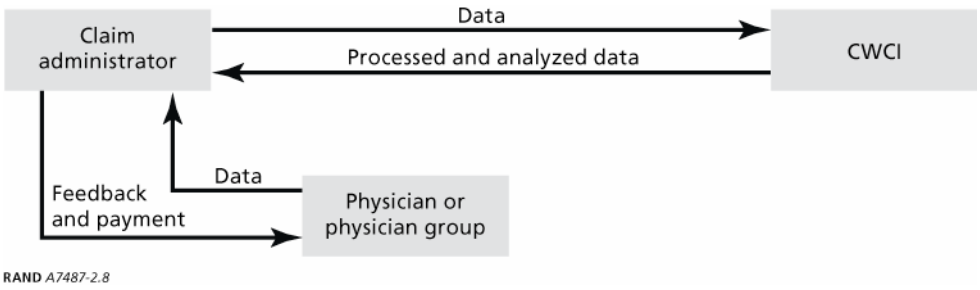
Figure 2.7. Data Collection and Evaluation Are Through an Independent Organization



Option 4: CWCI Collects and Evaluates the Data

This model (Figure 2.8) would take advantage of CWCI's current data infrastructure but may exclude payers who do not belong to CWCI from participating. It would be best suited for a voluntary system that does not involve public reporting, because the full complement of care furnished by a provider to injured workers would not be captured. Given the high level of distrust within the WC system, it is less likely than options 2 or 3 to build the provider confidence in the measurement system needed for successful operation.

Figure 2.8. Data Collection and Evaluation Are Through CWCI



Participants at the WC P4P roundtable suggested that, to build trust in the program, the data collection and assessment processes need to be transparent, the data need to be audited, and an independent evaluation of the program's effects need to be conducted.

HOW SHOULD THE REWARD BE STRUCTURED?

P4P programs strive to structure rewards that are of sufficient magnitude to capture physicians' attention, establish incentives for as

many physicians as possible, have an easily understood incentive formula, and can be awarded in a timely manner (Sorbero, Damberg, et al., 2006). The development of the reward structure includes several decisions. Key issues that must be determined include

- the structure of the financial reward
- the criteria for receiving a reward
- the mechanism through which the reward is financed
- whether nonfinancial rewards should be included.

Form of the Financial Reward

The financial reward can be constructed as a lump-sum bonus or a modified fee-schedule payment. The latter could include either increasing payment for already reimbursed activities or providing explicit payment for specific work-related activities that are not separately paid under OMFS (such as completion of certain forms).

Option 1: Bonuses

Lump-sum bonuses could be administered annually based on a physician's performance during the prior year. Bonuses have the advantage of being explicit and, if sufficiently large, may be effective in capturing a physician's attention and engagement in the program. The total amount allocated toward the bonus pool can also be set in advance; total payouts stay within the total amount set aside, providing certainty in budgeting. This approach, compared to the alternatives, is also relatively easy to administer.

Option 2: Modified Fee-Schedule Payment

The fee schedule might be modified in several ways:

- Pay for specific work-related activities that are not currently explicitly paid for under OMFS. In this regard, proposals have been made to create *Current Procedural Terminology* (CPT), 4th edition (CPT-4) codes that describe return-to-work–management activities performed by primary treating physicians that are in addition to medical management of the case. For example, the Washington demonstration program paid providers for calling employers (Wickizer et al., 2004).

- Pay for compliance with specific administrative processes. For example, the Washington demonstration paid for timely filing of complete first reports of injuries (Wickizer et al., 2004); higher payments for electronic billing or for reporting quality measures is another example that might be considered.
- Establish higher levels of fee-schedule payments to physicians based on performance in the prior year.

Any of these strategies could be used by payers in a voluntary P4P system, because the current OMFS is not binding if the payer and provider agree to an alternative level. For medical networks, employers and payers are allowed to establish discounted fees but are prohibited from establishing any payment incentives that are intended to reduce, delay, or deny medical treatment or restrict access to care. Modifications of OMFS to incorporate these types of rewards would be tantamount to creating a mandatory P4P (of which a payer and provider could agree to opt out). Incorporating the additional allowances into OMFS may be the most efficient and effective way to align incentives and overcome the distrust between some payers and providers.

The anticipated revision in the physician fee schedule is an opportune time to consider incorporating some P4P rewards into OMFS. A study by The Lewin Group on physician effort with injured workers concluded that evaluation and management (E/M) services provided to injured workers require about 28 percent more effort than services to Medicare patients and recommended that the fee-schedule rates be adjusted accordingly (Dobson et. al., 2002). Recently, the DWC administrative director raised OMFS payments for E/M services to Medicare levels. In lieu of additional, across-the-board increases for these services when the new physician fee schedule is implemented, it may be more effective to create payments for specific work-related activities.

A substantial drawback of the modified fee-schedule approach for specific activities is that the payer would not know in advance what the total liability would be and thus how to establish a "fixed" budget for incentive payments. The total amount paid in any year would vary based on physician compliance with the measures. Also, a relatively small

increase on specific services may be less likely to capture providers' attention than a lump-sum bonus would, even if the magnitude of the incentive for a year is the same dollar amount. However, the experiences of the Washington pilot project (Wickizer et al., 2004, 2007) suggest that this may not be the case. To avoid creating an incentive to increase the use of all services, the fee-schedule payment should only be increased specific services that the payer wants performed and is measuring.

Criteria for Receiving the Reward

There are four core options available to serve as the basis for determining whether a physician is eligible to receive an incentive payment. These options can be used alone or in combination. The determination is made based on

1. absolute threshold of performance or performance target (e.g., activity prescription is documented in 95 percent of patient evaluations).
2. relative threshold of performance (e.g., above the seventy-fifth percentile of all providers' or peer groups' scores for activity prescription).
3. action or service provided (e.g., additional payment [or a point system] each time the physician documents on a patient-evaluation bill that an activity prescription was provided).
4. improvement in performance over time (e.g., the physician's rate of prescribing activity increases from 70 percent to 85 percent of patient evaluations).

Option 1: Absolute Performance Threshold

This approach would require determining in advance for each performance measure what level of performance would result in payment. In the absence of empirical data on which to base determinations or a priori knowledge about the appropriate use of a service (i.e., to account for variations in practice and allow for patient refusal or those not visiting the physician to mitigate incentives to not treat nonadherent patients), the determination of absolute performance thresholds can be challenging.

A key strength of absolute performance thresholds is that they provide clear messages to physicians about what level of performance must be achieved to receive an incentive. They also help physicians understand the likelihood of getting an incentive. A potential problem is that, when incentives go toward reaching a common, fixed, performance target, little benefit is achieved, since the incentive motivates physicians only to meet the target and not go beyond it (Dudley et al., 2004; Rosenthal et al., 2005) and poorer performers who do not expect to see a reward may be discouraged from investing in quality improvement. This could be addressed by having a sliding scale of absolute targets or by adjusting targets upward on a periodic basis.

Use of absolute performance targets can make budgeting for the incentive more challenging. It is not known in advance how many providers will reach the threshold and thus will have to be paid. This problem can be addressed by setting a fixed budget for each measure. Then the actual payout to each individual provider will depend on the number of providers meeting the threshold. This will, however, create uncertainty among physicians about the award amount they could expect to see if they hit the performance target.

Option 2: Relative Thresholds

An example of a relative threshold is pegging the reward to the seventy-fifth percentile performance rate of all providers in a performance pool. Relative thresholds have two advantages: (1) the absolute level of the performance threshold continues to increase as overall performance by providers improves and (2) it is easier to budget for this type of threshold than for an absolute threshold, since the number of providers receiving the incentive is known a priori (e.g., if the reward is pegged to the seventy-fifth percentile, then those in the seventy-fifth percentile or higher will be paid; so the number of physicians that will fall in the top quartile of performance can be calculated). The disadvantages with relative thresholds are that (1) providers will not know in advance how well they must perform to receive the incentive, so it increases a physician's uncertainty about the likelihood of receiving a reward; (2) the absolute value of the threshold will be arbitrary and may not reflect superior performance

(e.g., if performance scores are uniformly low on a particular measure, the ninetieth-percentile performance might be 40 percent); and (3) the differences in performance between providers who receive and do not receive the incentive may be very small.

Both absolute and relative targets result in rewarding the physicians who were already providing high quality of care prior to the start of the program (Dudley et al., 2004; Rosenthal et al., 2005). In consequence, poorer performers, not expecting to see a reward, may be discouraged from investing in quality improvement.

Option 3: Payment for Provision of a Service

This method is essentially enhanced FFS on a specific-service basis. For example, rather than measuring rates of a service, the payment for the service would be increased. The advantages of this approach are that it allows providers to know what must be done to obtain the incentive and that it removes the necessity for providers to have adequate numbers of patients to create rates for individual measures. Furthermore, a recent commentary by P4P experts suggests that this is a very effective approach (Rosenthal and Dudley, 2007). The disadvantage is that budgeting for the higher payments can be challenging, because it is not known in advance the extent to which providers will respond to the enhanced payment.

Option 4: Improvement in Performance

This approach would pay all physicians who achieve improvements from one year to the next or relative to a baseline measure. This strategy benefits physicians whose performance is relatively low by allowing them to reap some of the rewards; as a consequence, it may engage a larger number of physicians in the program than do methods that reward only high performers. This approach penalizes, or may at least discourage, providers who are already doing well and thus have less room to improve (Rosenthal et al., 2005). One challenge to implementing this approach is that it is not unusual for performance-measurement specifications to change year to year, which may make year-to-year comparisons of performance challenging depending on the nature and extent of the changes.

The choice of criteria to use for determining the reward will depend in part on the level of accountability. If it is the individual physician, then payment for the provision of a service will likely be the only feasible approach due to the small number of WC patients cared for by individual physicians. The same may be true at the medical-group level, depending on the size of the group. Focusing measurement at the level of the network would allow consideration of other criteria for determining reward.

Mechanism for Financing Rewards

The underlying assumption of a WC P4P program is that the financial incentives will foster value-based medical care and should reduce total WC system costs in the long run. This is supported by the evaluation of the Washington pilot project that showed that saving increased over time (Wickizer et al., 2007). Bearing in mind that this assumption is largely untested, there are several ways that monetary awards could be financed. The first way would be to cover the costs of the financial rewards though employer contributions (either as self-insured employers or through WC insurance premiums). Employers or payers would assume the risk that the program will generate system savings, and physicians would not bear any costs for the financial rewards. The second way would be to cover the costs of the financial rewards through modifications in the physician fee schedule, such as reduced updates for inflation, reductions in payments for particular services, or withholdings in fee-schedule payments. If aggregate physician payments remained the same but were redistributed among physicians based on performance, physicians would bear the cost of the monetary rewards and employers and payers would retain any system savings created by the program. The third alternative, a shared-saving approach, would base provider bonus payment amounts on the savings generated by the improved performance, so employers or payers and providers would share the costs and benefits for operating the program. The shared-saving approach is more difficult to implement than the other alternatives because it requires timely and credible measures of savings attributable to performance.

Option 1: Finance Through Employer or Payer Contributions

Under this approach, rewards would be financed through insurance premiums and self-insured employer expenditures; that is, the additional payments to providers to reward performance would be treated as a WC medical expense. The actual impact on employer expenditures would ultimately be determined by the extent to which the financial incentives are effective in fostering value-based care and generating overall system savings.

This approach may be the only feasible way to implement a P4P program in California's current WC environment for several reasons. First, current payments to physicians and other practitioners have been constrained by relatively low maximum allowable fees under OMFS, the lack of regular updates for inflation, the 5-percent reduction required by California Senate Bill 228, and the fee discounting required by some medical provider networks. Further reducing payments to some physicians to reward others could create access problems. Second, there is considerable contention in the system, and providers may not have sufficient trust that the P4P program will be administered fairly to give up any portion of their payments in expectation of being rewarded for improved performance.

Option 2: Finance by Redistributing Physician Payments

Another way to finance the reward system would be through reducing payments to lower-performing physicians and increasing payments to higher-performing physicians while holding aggregate payments constant. There are several ways this could be done:

- A withholding could be imposed on all participating physicians that would be paid out as a bonus to high performers at the end of the year. For example, a 2-percent reduction could be made in fee-schedule payments that would accumulate in a pool that would be distributed at year end to reward performance. Those physicians who are not eligible for a reward would experience the 2-percent payment reduction. This approach could work under either a mandatory or voluntary program.
- To avoid reducing physician fees, some P4P plans finance rewards by putting aside the portion of a regularly scheduled

fee increase. While OMFS does not currently have regular updates, this approach may be an option under the new physician fee schedule if, for example, annual increases for inflation are tied to the increase in the Medicare Economic Index. Its advantage over a withholding is that no physician experiences a fee-schedule reduction; those who are not eligible for a reward receive a smaller increase than that to which they would be entitled in the absence of the P4P program. This approach could work under either a mandatory or voluntary program.

- The new physician fee schedule provides another potential way to finance the rewards of a mandatory program. The administrative director of DWC has considerable latitude in how the new physician fee schedule is structured and in determining aggregate payment levels. Financial rewards that are paid through OMFS could be taken into account in establishing new maximum allowable fees. The rewards could range from explicit allowable fees for certain work-related components of E/M services such as activity prescription, report writing, and communication with employers to across-the-board higher maximum allowable fees for physicians meeting certain performance criteria. If the fee-schedule rewards are "cost neutral," they would be financed through lower payments for other services.

Option 3: Finance Through Shared Savings

Under a shared-saving approach, providers receive rewards only if the P4P program as a whole demonstrates savings to the WC system over projected costs without the program in place. The estimated savings are then shared between the program sponsor and providers. This approach is the most complex to implement, because it requires a credible and timely determination of savings attributable to the P4P program. It is rarely used by group health P4P programs because of the difficulties of determining the savings on a timely enough basis for the rewards to affect behavior. At best, it may be feasible under a voluntary, single-payer WC program involving providers with large WC patient loads. Care

would need to be taken to structure the performance measures and rewards in such a way that they would not violate the labor code's medical provider network provision that prohibits structuring physician compensation "to achieve the goal of reducing, delaying, or denying medical treatment or restricting access to medical treatment" (California Labor Code §4616. One way to address this might be through tying the reward to savings in nonmedical expenses, such as reductions in temporary-disability payments.

In general, participants at the WC P4P roundtable discussion felt strongly that a "carrot" rather than a "stick" financial incentive would encourage provider buy-in within the current environment. This would suggest that the rewards should be financed by payer or employer contributions.

Use of Nonfinancial Rewards

Not all incentives need to be financial rewards directed at the health-care provider. Nonfinancial incentives can also be used to motivate provider behavior. Here, we discuss two such mechanisms: public reporting and reductions in administrative hurdles.

Option 1: Public Reporting

Public reporting is a strategy in which health plans or other entities provide consumers with comparative information on specific health-care providers and their peers. For example, in the California P4P program sponsored by IHA, OPA creates an annual, consumer-focused public score card for medical-group performance. The score card is based on IHA-supplied, aggregated P4P data that summarize whether patients are getting appropriate care and data from the Consumer Assessment Survey (CAS) used to measure patient experience and satisfaction with care.

The Ohio Bureau of Workers' Compensation produces an annual report card to inform employers insured through the state fund about the managed-care organizations that help manage WC claims and coordinate medical services. The report card has measures for the timeliness of first reports of injury processes, optimal return-to-work experience,

and injured-worker and employer satisfaction with managed-care organizations (Ohio Bureau of Workers' Compensation, undated).

Given the lack of clinical-performance data for WC conditions, public reporting is most feasible in the short run to measure satisfaction with care. It could be structured to provide injured workers with information on the medical groups available within a medical provider network or to provider employers and payers with information at the medical provider network level. A clinical component could be added as clinical-performance measures are developed and the P4P program matures. Appropriate risk adjustment would be needed so as to not penalize physicians with more complex patient loads.

Option 2: Reduce Administrative Hurdles

As stated previously, the WC program involves more administrative processes and reporting requirements than nonoccupational health-insurance programs. A potential nonfinancial reward would be a reduction in the administrative burden on high-performing providers. Utilization management is one area in which this might occur. Through interviews with individuals involved with utilization management in the California WC program, we learned that most payers currently require all providers to obtain prospective authorization for medical treatments but are seeking ways to refine their UR processes and eliminate unnecessary administrative burden. A targeted UR approach could reward physicians whose patterns of care are consistent with American College of Occupational and Environmental Medicine (ACOEM) guidelines. Provider profiling—at the medical-group or network level—could be used to identify those physicians whose denial rates fall below some defined threshold for designated procedures. These physicians could be either be exempted from prospective review or reviewed on a more limited basis, reducing administrative burden for providers and payers alike (Nuckols et al., 2005). Washington WC administrators have implemented a program to reduce URs for providers practicing within state Department of Labor and Industries' Medical Treatment Guidelines (Washington State Department of Labor and Industries, undated).

WC stakeholders were enthusiastic about the inclusion of nonfinancial rewards in a WC system. In particular, reducing the administrative burden on providers in the UR process was seen as a substantial benefit.

SUMMARY

This chapter describes a host of design components and options that will need to be considered in designing a P4P program for WC. The current published literature reveals an absence of scientific evidence to suggest a single, best strategy for designing a P4P program that is likely to yield maximum benefits—whether the goal is to improve quality or reduce costs or do some combination of the two. There is little empirical knowledge about the comparative impact of design-component options (e.g., basis for attributing care to a physician, differential impact of using performance thresholds or year-to-year improvement, few versus many measures) or about the circumstances under which the various design components are more or less likely to have an impact (i.e., local market characteristics, such as organization of physician practices and exposure to private P4P programs sponsored by individual organizations or regional coalitions).

Much of what we know about P4P program design comes from the early experiences of P4P program sponsors, who are learning step by step, in trial-and-error fashion, and modifying their programs as they go. Participants at the WC P4P roundtable had valuable insights into the types of structures that would be acceptable to providers and payers. Due to the lack of experience with P4P in WC, it will be critical to pilot-test all aspects of a new program, be it a demonstration or a full-scale program implementation. This will allow providers to gain experience with the new program, as well as facilitate opening the lines of communication in a currently distrustful environment. Generally, the program's goals and objectives will guide the selection of the performance measures and the decisions on the reward structure. However, the choice of performance measures, as well as the decisions on the data-collection and performance-evaluation processes, will also be

influenced by the availability of appropriate data, measures, and infrastructure to support performance assessment.

CHAPTER THREE. POSSIBLE PAY-FOR-PERFORMANCE MODELS FOR WORKERS' COMPENSATION

This chapter synthesizes the various options discussed in Chapter Two into three models that might be considered for a collective P4P initiative within the California WC system. Discussion focuses on two models for a mandatory program implemented through OMFS and a collective, voluntary program among payers. We do not discuss a voluntary, individual-payer model, because payers can negotiate P4P contractual arrangements with providers without additional policy development and arrangement particulars will be specific to the payer and providers. Most design options discussed in Chapter Two would also be relevant to an individual-payer P4P program. These programs may be the most feasible in the short run and could serve as valuable pilot tests if they incorporated an evaluation component.

Before we describe the models, we first discuss an overall approach we view as necessary for a P4P program to succeed. Next, we address the main design components for P4P program models, which were presented in Chapter Two: goals and objectives, performance measures, reward structure, data collection and evaluation, and financing for each of the models. We describe suggestions for what might be the appropriate next steps in evaluating the advisability and feasibility of each model.

Overall Approach

For a P4P program to succeed, particularly in an environment with a relatively high level of distrust, all parties involved would need to see it as a "win-win." The best way to ensure that this occurs is to involve relevant stakeholders in the design of a P4P program. Possibilities for mutually beneficial approaches include establishing safeguards and processes to build trust among stakeholders. Selecting performance measures that are anticipated to generate overall savings through improved quality and better work-related outcomes would result in a program expected to benefit not only the physician, but the

employer as well. Starting small with a pilot program would enable program developers to determine realistic goals, test measures, assess the reporting burden associated with the program, and determine the associated costs and savings. In an environment of distrust and physicians feeling that they are not adequately paid on a FFS basis, it is important to create financial incentives that provide financial rewards for good performance without reducing payments for other providers. Otherwise, providers will feel that the program is simply payers' effort to cut costs. Providers are also likely to be responsive to nonfinancial rewards such as reductions in UR requirements. Paying for additional reporting needed to measure quality of care would also increase provider receptivity to the program.

The goals and objectives of a P4P program could be independent of whether a program has mandatory or voluntary participation. Goals most likely to be supported by a variety of WC stakeholders include improvement in the quality of care delivered in the WC system, increased patient satisfaction, and improved work-related outcomes. Improved efficiency in delivery method is likely to be supported by both payers and employers but is less likely to be supported by physicians unless they can be provided specific examples of waste in the system. The identified goals should remain in the forefront to guide the development of the rest of a P4P program.

A variety of roadblocks would have to be considered in designing the program. These include the complexity of the current system, the lack of clinical-performance measures, the absence of a developed central data-collection system and ongoing performance monitoring system to facilitate evaluation, and the involvement of multiple payers and many physicians who treat only a few injured workers each year. For this reason, the models start with relatively straightforward measurement and reward structures but could be enhanced over time as the measurement and data-collection systems mature.

Table 3.1 presents concepts for measures that WC roundtable participants recommended. These concepts represent the first step in defining P4P program goals and in developing appropriate performance measures. Additional work would be needed to define measure

specifications that could be used in a P4P program. In particular, a time-consuming and expensive multistep process is required to develop measures of clinical performance. Treatment guidelines exist in WC for a number of clinical areas that can provide a framework on which clinical measures are based. However, guidelines provide general guidance, while measures are very specific and explicit. The steps involved in developing measures include (1) identifying the clinical topics to address, (2) developing preliminary indicators, (3) reviewing the literature for each preliminary indicator, (4) formally evaluating indicators by expert panels, (5) developing a data-application tool, and (6) pilot-testing and refining final indicators.

Table 3.1. Potential Measures for a Workers' Compensation P4P Program

Clinical	Efficiency	Patient Experience	Administrative	Work-Related Outcomes
Referrals consistent with guidelines Repeat surgeries to same body part	Total costs (combined medical and temporary total disability)	Retention rates	Cases litigated Work and activity assessment completion rate Modified duty rate Noncertification of referral rate Completion of training on WC system	Return-to-work rate Number of days at temporary, total disability Time to return to work

Building on existing data infrastructure and reporting systems would reduce data-collection costs for providers and minimize administrative costs for payers. Ideally, the codes needed to measure performance would be incorporated into the billing system so that costly chart review and medical-record abstraction would be unnecessary. The code set used for P4P should conform to the administrative-simplification requirements of the Health Insurance Portability and Accountability Act of 1996 (HIPAA) (Public Law 104-191). Under the HIPAA provisions, the U.S. Department of Health and Human Services has

established national standards for electronic health-care transactions, including the medical code sets used to identify what procedures, services, and diagnoses pertain to a patient encounter.[3]

The HIPAA-approved medical codes for physician services are a combination of CPT, maintained by the AMA, and the Healthcare Common Procedural Coding System, maintained by Medicare for services that are not described in CPT. If the P4P program needs codes that are not in these code sets, California stakeholders would need to work with the AMA's CPT editorial panel and other interested parties such as other state WC programs and ACOEM to have them incorporated into CPT.

Depending on medical activities selected for additional payment, additional codes may not be required. For example, Washington's P4P program, described in Chapter Two, reimbursed injured workers for calls to their employers to coordinate returning to work (Wickizer et al., 2004) and rewarded physicians who filed timely reports. Codes are not needed to measure the timeliness of report filing. CPT has codes for telephone calls that might be considered if this activity were to become separately payable. (Neither the current OMFS nor the Medicare fee schedule has relative values for telephone calls.) There are also specific CPT codes for care coordination. New codes would most likely be needed to describe work-related disability-management activities and to measure clinical performance. Two levels of codes might be considered:

- Category 1 CPT codes describe most physician procedures and services. To establish a new category 1 CPT code, the CPT editorial panel requires that many physicians or other health-care professionals perform the service or procedure across the country in multiple locations and that the service or procedure's clinical efficacy be well established.

[3] Many physicians who provide care to injured workers submit paper bills via mail or fax. These physicians are not required to use the HIPAA-approved code set. However, we believe that any P4P program should encourage electronic billing and conform to the HIPAA standards for electronic transactions.

- Category 2 CPT codes are optional tracking codes created for performance measurement. They are intended to facilitate data collection by coding certain services and tests that improve quality or accountability, including specific activities that are typically included in an established (E/M) visit. A performance-measurement advisory group initially reviews requests to evaluate whether the qualifying criteria are met, including the extent to which an evidence-based process was used to develop the measure, a multidisciplinary review process was used to develop consensus on the measures among all constituents, and the measures were tested for validity and feasibility (AMA, 2007).

MODEL 1: MANDATORY PROGRAM PAYING ON FEE-FOR-SERVICE BASIS

A mandatory P4P program could take a relatively simple form that would require minimal changes to the existing system. We describe the system as *mandatory* because it would be implemented by establishing fees through OMFS for specific services. Payers would be required to pay these fees unless they negotiated an alternative arrangement with the affected physicians.

Reward Structure

OMFS could be modified to include explicit fees for disability prevention and management activities that are unique to work-related injuries. This approach of paying providers an additional fee for each desirable activity has been noted to create the incentive for every provider to deliver the best care to each patient rather than just rewarding those providers identified as top performers (Rosenthal and Dudley, 2007).

Performance Measures

Explicit payment for selected work-related activities on a per-service basis has two advantages over other types of rewards. First, it does not require the development of performance measures—that is, the development of measure specification detailing the types of patients that fall into the numerator and denominator. Instead, it requires the

definition of a set of codes and documentation requirements for activities to be newly reimbursed or codes for which the fee-schedule amount would be increased. Second, it eliminates the need for adequate patient volumes to construct measures. As long as a physician treats a single WC patient, he or she can be rewarded for engaging in the desired activity that promotes high-quality care.

Data Collection and Evaluation

Paying for specific services on an FFS basis requires no changes in the current data structure. WCIS could be used to measure the extent to which services are actually provided and the aggregate payments for the services.

Financing

The additional payments would be classified as medical expenses and assumed by employers or payers. Presumably, the wide-scale P4P program implementation would occur only after the activities were determined to be cost-effective in a pilot program.

EXAMPLE 2: MANDATORY PROGRAM THAT REWARDS OVERALL PERFORMANCE

A more complex, mandatory P4P program would add more administrative processes to the system but would facilitate the development of a broader set of performance measures over time. The system would be mandatory for physicians: All physicians relevant to the selected measures would have their performance measured automatically and would not be given the choice to participate. Data for performance measurement would be collected through WCIS, and uniformly determined rewards would apply to high performers. Payers would be required to pay the resulting financial reward unless they negotiated an alternative arrangement with the affected physicians. Unlike the simpler FFS model, this structure would provide an opportunity for pooling data and providing report cards at the medical-group or network level when there are enough observations.

Reward Structure

DWC, in consultation with CHSWC and the various constituents in the P4P program, would need to decide on the reward structure. The most feasible structure in the short run would be a two-tier fee schedule. High-performing physicians (i.e., gold-star physicians) in the measurement year would receive higher fee-schedule payments than would low-performing physicians in the succeeding payment year (e.g., a 2-percent add-on).[4] Individual payers might further reward designated high-performing physicians through the use of nonfinancial incentives, such as reduced UR.

Performance Measures

In the absence of clinical measures for WC, an initial set of process or structural measures that would apply broadly to physician specialties treating injured workers might be considered. The focus in this example would be on primary treating physicians who provide care to a significant number of injured workers. The following criteria are illustrative of those that might be considered in determining eligibility for gold-star designation and higher payments:

- Treat a minimum number of injured workers in the clinical practice.
- Complete a continuing-education course on caring for injured workers and disability management.
- Regularly submit data abstracted from medical charts to facilitate construction of clinical measures and risk adjustment.
- Document activity prescription during a high percentage of medical encounters.

[4]A bonus for current-year performance is a more attractive policy because it directly ties the bonus to performance in the same year and aggregate bonus payments can be determined in advance; however, the multiple payers involved in WC pose a number of administrative hurdles to administering a single bonus pool that make higher payments in the succeeding year the more feasible approach.

- Document communication with employers in a high percentage of claims.
- Experience a high patient-retention rate.

The above measures could be developed from WCIS data once the system is fully implemented and should not require risk adjustment. Patient-satisfaction measures would be desirable if public reporting were incorporated into the reward structure, but these measures would require separate data collection. Quality measures could be added as they are developed.

Data Collection and Evaluation

DWC should administer a mandatory program. However, an external organization could support DWC in the effort, facilitating the development and testing of performance measures and the annual evaluation of overall performance using WCIS data.

Financing

The additional payments to high-performing physicians would be classified as medical expenses and assumed by employers and payers. As was the case with the other mandatory model, the cost-effectiveness of a pilot program should be evaluated before wide-scale implementation of the P4P program occurs.

MODEL 3: COLLECTIVE VOLUNTARY PROGRAM WITH PAYER-DETERMINED FINANCIAL REWARDS

A voluntary program could be modeled after the IHA P4P program, in which multiple payers voluntarily join together to create a collaborative. The payers involved in the collaborative would use a standardized set of core measures, with individual payers having the option to use additional measures as desired. Data for the core set of measures would be pooled to increase sample size for individual physicians.

Reward Structure

Individual payers would have the option to decide how to create incentives for physicians. Incentives could take the form of a bonus payment to high-performing providers or those who have shown substantive improvement in their performance over time. In addition, a component of pay-for-reporting—paying physicians for submitting data abstracted from medical charts—could be included to offset the costs of collecting additional data beyond what is currently captured in administrative data that may be needed to construct measures.

Not all rewards in a P4P program need to be financial. Nonfinancial incentives also can be effective motivators of behavior change. Nonfinancial incentives could involve identifying a group of high-performing providers that would experience fewer administrative burdens for providing care in the WC system, such as reduction in UR or in requirements of precertification of referrals. Another reward could be prompt payment of submitted bills. Currently, physicians are supposed to be paid within 15 days of the payer receiving a complete bill from the physician, but providers at least *perceive* that payers make payments more slowly. Finally, public reporting of performance scores is a frequently used and effective method to stimulate quality improvement (Hibbard, Stockard, and Tusler, 2003, 2005).

Performance Measures

Performance measurement for a P4P program could be based on one or more expensive conditions such as lower-back pain or carpal-tunnel syndrome. NCQA has established measures for lower-back pain, and RAND researchers are currently developing clinical-process measures for carpal-tunnel syndrome that are expected to be tested and validated in approximately a year. Additional measures would need to be developed if additional clinical areas are included.

Data Collection and Evaluation

Multiple approaches to data collection and evaluation could work in the context of a voluntary P4P program.

- An independent organization could pool and process the data that the payers provide. This organization would be responsible for

providing feedback reports to both participating payers and the providers in their networks on the core set of measures using the pooled data.

- CWCI could collect and evaluate the data. CWCI currently receives data from some payers and self-insured employers that voluntarily submit it for research purposes. The infrastructure that CWCI has in place makes it a natural entity to take on the responsibility of data collection and evaluation for P4P. However, this approach would exclude organizations that do not belong to CWCI from participating in a voluntary P4P program or force them to affiliate with CWCI.
- The P4P program could build upon the current DWC-maintained data flows for data-analysis and performance feedback. DWC could either analyze and evaluate the data itself or contract with an independent third-party organization to evaluate the data.

Financing

Each employer or payer would have discretion to determine how the reward structure would be financed. Participating physicians would need to agree to the extent the financing arrangements would alter the amounts payable under OMFS.

NEXT STEPS

As noted throughout this paper, there are challenges to implementing a WC P4P, including the lack of clinical measures for WC conditions, multiple payers, and the many physicians who treat only a few WC patients. This paper offers three models that we believe might be able to surmount these problems, provided that the stakeholders have the commitment and trust to work through the design issues and allow the P4P program to evolve over time. Given the current WC environment and amount of change that has occurred in the medical treatment system over the past few years, the various stakeholders need to confirm their willingness to undertake a collective P4P initiative. Thus, a critical next step is for CHSWC to expand the discussion to include representatives of the various stakeholder constituencies to gauge their

levels of interest and commitment in a P4P initiative, define the program's goals and objectives, and identify potential "idea champions" to promote the P4P concepts within the California WC stakeholder community. If there is sufficient interest, a structure could then be established for further collaborative work on design options and issues. If there is not interest in a collective effort, there may nevertheless be interest in pilot projects to ascertain whether a P4P program would be cost-effective. Further, there may be interest in workshops and joint educational activities on ways that individual employers or payers could incorporate P4P principles into dealings with physicians.

While the current California WC environment poses challenges to advancing the P4P concepts, DWC has the opportunity to put mechanisms in place to measure and reward physician performance. The opportunities include the following:

- Data collection has been initiated for WCIS, and the system should become operational by the end of 2007. This is an opportune time to consider how the medical data collected through WCIS can be turned into useful information to support performance evaluation. The critical decision is whether WCIS will support monitoring and evaluation at the system level only or whether it will be structured to also support measures of performance at the individual, group, or medical provider network level.
- DWC has begun work to modify the physician fee schedule. As the goals for the new fee-schedule structure are established, consideration should be given to how to align the financial incentives inherent in the fee schedule with value-based care.

P4P alone will not be sufficient to drive value-based medical care provided to injured workers; rather, it should be considered as part of a multipronged set of strategies designed to increase the efficient delivery of high-quality care that enables rapid and sustained return to work. Other strategies recommended in Wynn (forthcoming) examining the medical treatment system are to

- establish an ongoing monitoring system to assess system performance: access, quality, cost, utilization, and patient satisfaction.
- develop clinical criteria to measure appropriate care.
- make evidence-based treatment guidelines on common WC conditions and modalities readily available.
- adopt a new physician fee schedule.

Integration of P4P incentives with these strategies may accelerate the drive for value-based medical care in the California WC system.

REFERENCES

American Accreditation HealthCare Commission/URAC, *Measuring Quality in Workers' Compensation Managed Care Organizations: Technical Manual of Performance Measures*, Washington, D.C.: URAC, 2000. As of August 7, 2007: http://www.urac.org/savedfiles/WorkersCompPerformanceMeasures.pdf

American Medical Association, *Current Procedural Terminology*, 4th ed., Chicago, Ill.: American Medical Association, 2004.

———, "CPT® Process," last updated April 30, 2007. As of August 7, 2007: http://www.ama-assn.org/ama/pub/category/3112.html

Bosworth, Hayden B., Tara Dudley, Maren K. Olsen, Corrine I. Voils, Benjamin Powers, Mary K. Goldstein, and Eugene Z. Oddone, "Racial Differences in Blood Pressure Control: Potential Explanatory Factors," *The American Journal of Medicine*, Vol. 119, No. 1, January 2006, pp. 70.e9–70.e15.

Bridges to Excellence, "Spine Care Link," 2007. As of July 27, 2007: http://www.bridgestoexcellence.org/programs/scl.mspx

BTE—*see* Bridges to Excellence.

California Commission on Health and Safety and Workers' Compensation, *2006 Annual Report*, December 2006. As of July 30, 2007: http://www.dir.ca.gov/CHSWC/Reports/AnnualReport2006.pdf

California Division of Workers' Compensation, "Official Medical Fee Schedule," May 2007a. As of August 1, 2007: http://www.dir.ca.gov/dwc/OMFS9904.htm

———, "Workers' Compensation Information System (WCIS)," July 2007. As of August 1, 2007: http://www.dir.ca.gov/dwc/WCIS.htm

California Division of Workers' Compensation, and Bickmore Risk Services, *A Study of the Effects of Legislative Reforms on California Workers' Compensation Insurance Rates*, California: DWC, March 2006. As of August 1, 2007: http://www.dir.ca.gov/dwc/Study_legislativeReformsCaWCInsuranceRates/Study_legislativeReformsCaWCInsuranceRates.html

California Office of the Patient Advocate, "2006 Healthcare Quality Report Card: Rating California's HMOs and Medical Groups," Web page, 2006. As of August 1, 2007: http://opa.ca.gov/report_card/medicalgroupcounty.aspx

California Senate Bill 228, Workers' Compensation, September 30, 2003. As of August 7, 2007: http://www.leginfo.ca.gov/pub/03-04/bill/sen/sb_0201-0250/sb_228_bill_20031001_chaptered.pdf

CHSWC—*see* California Commission on Health and Safety and Workers' Compensation.

DiMatteo, M. R., "Variations in Patients' Adherence to Medical Recommendations: A Quantitative Review of 50 Years of Research," *Medical Care*, Vol. 42, No. 3, 2004, pp. 200–209.

Dobson, Al, Joan DaVanzo, Maria Consunji, and Jawaria Gilani, *A Study of the Relative Work Content of Evaluation and Management Codes*, Falls Church, Va.: The Lewin Group, April 29, 2003. As of August 1, 2007: http://www.dir.ca.gov/IMC/StudyRelWorkContEvalMmgtCodes(draft).pdf

Dudley, R. Adams, Anne Frolich, David L. Robinowitz, Jason A. Talavera, Peter Broadhead, and Harold S. Luft, *Strategies to Support Quality-Based Purchasing: A Review of the Evidence*, Rockville, Md.: U.S. Department of Health and Human Services, Public Health Service, Agency for Healthcare Research and Quality, AHRQ publication 04-0057, July 2004. As of August 1, 2007: http://purl.access.gpo.gov/GPO/LPS71429

DWC—*see* California Division of Workers' Compensation.

Ellis, Randall P., and Thomas G. McGuire, "Hospital Response to Prospective Payment: Moral Hazard, Selection, and Practice-Style Effects," *Journal of Health Economics*, Vol. 15, No. 3, June 1996, pp. 257–277.

Farley, Donna O., "Relevance of the CAHPS® Consumer Assessment Survey for Workers' Compensation Medical Care," in Stephanie S. Teleki, Allard E. Dembe, Jeffrey S. Harris, Thomas Wickizer, Donna O. Farley, and Barbara O. Wynn, *Research Colloquium on Workers' Compensation Medical Benefit Delivery and Return to Work*, Santa Monica, Calif.: RAND Corporation, CF-214-ICJ, 2006, pp. 83–102. As of August 1, 2007: http://www.rand.org/pubs/conf_proceedings/CF214/

Frolich, Anne, Jason A. Talavera, Peter Broadhead, and R. Adams Dudley, "A Behavioral Model of Clinician Responses to Incentives to Improve Quality," *Health Policy*, Vol. 80, No. 1, January 2007, pp. 179–193.

"Generous Provider Incentives Deliver Dramatic Returns," *Disease Management Advisor*, Vol. 12, No. 4, 2006, pp. 40–44.

Hibbard, Judith H., Jean Stockard, and Martin Tusler, "Does Publicizing Hospital Performance Stimulate Quality Improvement Efforts?" *Health Affairs*, Vol. 22, No. 2, March–April 2003, pp. 84–94.

———, “Hospital Performance Reports: Impact on Quality, Market Share, and Reputation,” *Health Affairs*, Vol. 24, No. 4, July–August 2005, pp. 1150–1160.

Hill, E.G., “State Policy Approach: Promoting Health Information Technology in California.” Sacramento, CA: Legislative Analyst’s Office (LAO), February 2007. Accessed July 30, 2007 at http://www.lao.ca.gov/laoapp/main.aspx.

IHA—*see* Integrated Healthcare Association.

Integrated Healthcare Association, “Integrated Healthcare Association Announces New Efficiency Measure in Pay for Performance to Improve Healthcare Quality and Reduce Costs,” press release, Los Angeles, Calif., February 15, 2007. As of August 1, 2007: http://www.iha.org/021507.pdf

Institute of Medicine, *Crossing the Quality Chasm: A New Health System for the 21st Century*, Washington, D.C.: National Academy Press, 2001.

———, *Performance Measurement: Accelerating Improvement*, Washington, D.C.: National Academies Press, 2006.

IOM—*see* Institute of Medicine.

Kaplan, Robert C., Narendra C. Bhalodkar, Edward J. Brown, Jessica White, and David L. Brown, “Race, Ethnicity, and Sociocultural Characteristics Predict Noncompliance with Lipid-Lowering Medications,” *Preventive Medicine*, Vol. 39, No. 6, December 2004, pp. 1249–1255.

Kominski, Gerald F., Nadereh Pourat, Dylan H. Roby, and Meghan E. Cameron, *Access to Medical Treatment in the California Workers’ Compensation System, 2006*, Los Angeles, Calif.: UCLA Center for Health Policy Research, February 2007. As of August 1, 2007: http://www.dir.ca.gov/dwc/AccessMedTreatmentReport2006/AccessToMedicalTreatmentInCAWC2006.html

Levin-Scherz, Jeffrey, Nicole DeVita, and Justin Timbie, “Impact of Pay-for-Performance Contracts and Network Registry on Diabetes and Asthma HEDIS® Measures in an Integrated Delivery Network,” *Medical Care Research and Review*, Vol. 63, No. 1, Supplement, February 2006, pp. 14S–28S.

Mattke, Soeren, Aruna Balakrishnan, Giacomo Bergamo, and Sydne J. Newberry, “A Review of Methods to Measure Health-Related Productivity Loss,” *The American Journal of Managed Care*, Vol. 13, No. 4, April 2007, pp. 211–217.

Medicare Payment Advisory Commission, “Public Meeting,” transcript, Washington, D.C., September 8, 2005. As of August 1, 2007: http://www.medpac.gov/transcripts/0905_allcombined_transc.pdf

MedPAC—*see* Medicare Payment Advisory Commission.

National Committee for Quality Assurance, “Back Pain Recognition Program,” 2007. As of July 27, 2007: http://web.ncqa.org/tabid/137/Default.aspx

NCQA—*see* National Committee for Quality Assurance.

Newhouse, Joseph P., “Do Unprofitable Patients Face Access Problems?” *Health Care Financing Review*, Vol. 11, No. 2, Winter 1989, pp. 33–42.

Nuckols, Teryl K., Steven M. Asch, M. A. Maggard, and Rebecca Shaw, *Quality of Medical Care in Workers’ Compensation: A California Demonstration Project. Study Design and Scope of Work*, unpublished RAND research, 2006.

Nuckols, Teryl K., Barbara O. Wynn, Yee-Wei Lim, Rebecca Shaw, Soeren Mattke, Thomas Wickizer, Philip Harber, Peggy Wallace, Steven M. Asch, Catherine MacLean, and Rena Hasenfeld Garland, *Evaluating Medical Treatment Guideline Sets for Injured Workers in California*, Santa Monica, Calif.: RAND Corporation, MG-400-ICJ, 2005. As of August 1, 2007: http://www.rand.org/pubs/monographs/MG400/index.html

Ohio Bureau of Workers’ Compensation, “MCO 2006 Report Card,” undated brochure. As of July 30, 2007: http://www.ohiobwc.com/downloads/brochureware/brochures/reportcard.pdf

OPA—*see* California Office of the Patient Advocate.

Petersen, Laura A., LeChauncy D. Woodard, Tracy Urech, Christina Daw, Supicha Sookanan, “Does Pay-for-Performance Improve the Quality of Health Care?” *Annals of Internal Medicine*, Vol. 145, No. 4, August 15, 2006, pp. 265–272.

Public Law 104-191, Health Insurance Portability and Accountability Act of 1996, August 21, 1996.

Public Law 109-432, Tax Relief and Health Care Act of 2006, December 20, 2006.

Rosenthal, Meredith B., and R. Adams Dudley, “Pay-for-Performance: Will the Latest Payment Trend Improve Care?” *The Journal of the American Medical Association*, Vol. 297, No. 7, February 21, 2007, pp. 740–744.

Rosenthal, Meredith B., Rushika Fernandopulle, HyunSook Ryu Song, and Bruce Landon, “Paying for Quality: Providers’ Incentives for Quality

Improvement," *Health Affairs*, Vol. 23, No. 2, March 1, 2004, pp. 127–141.

Rosenthal, Meredith B., and Richard G. Frank, "What Is the Empirical Basis for Paying for Quality in Health Care?" *Medical Care Research and Review*, Vol. 63, No. 2, April 2006, pp. 135–157.

Rosenthal, Meredith B., Richard G. Frank, Zhonghe Li, and Arnold M. Epstein, "Early Experience with Pay-for-Performance: From Concept to Practice, *The Journal of the American Medical Association*, Vol. 294, No. 14, October 12, 2005, pp. 1788–1793.

Schwarzenegger, Arnold, *Governor's Health Care Proposal*, January 8, 2007. As of August 1, 2007: http://gov.ca.gov/pdf/press/Governors_HC_Proposal.pdf

Shen, Yujing, "Selection Incentives in a Performance-Based Contracting System," *Health Services Research*, Vol. 38, No. 2, April 2003, pp. 535–552.

Sloan, Frank A., Derek S. Brown, Emily Streyer Carlisle, Gabriel A. Picone, and Paul P. Lee, "Monitoring Visual Status: Why Patients Do or Do Not Comply with Practice Guidelines," *Health Services Research*, Vol. 39, No. 5, October 2004, pp. 1429–1448.

Sorbero Melony E. S., Cheryl L. Damberg, Rebecca Shaw, Stephanie Telekie, Susan Lovejoy, Alison Dechristofaro, Jake Dembosky, and Cynthia Schuster, *Assessment of Pay-for-Performance Options for Medicare Phsician [sic] Services: Final Report*, Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, May 2006. As of August 1, 2007: http://aspe.hhs.gov/health/reports/06/physician/index.htm

Sorbero, Melony E. S., Andrew W. Dick, Jack Zwanziger, Dana Mukamel, and Nancy Weyl, "The Effect of Capitation on Switching Primary Care Physicians," *Health Services Research*, Vol. 38, No. 1, Part 1, February 2003, pp. 191–209.

Stover, Bert, Thomas M. Wickizer, Fred Zimmerman, Deborah Fulton-Kehoe, and Gary Franklin, "Prognostic Factors of Long-Term Disability in a Workers' Compensation System," *Journal of Occupational and Environmental Medicine*, Vol. 49, No. 1, January 2007, pp. 31–40.

Victor, Richard A., Peter S. Barth, and Te-chun Liu, *Outcomes for Injured Workers in California, Massachusetts, Pennsylvania, and Texas*, Cambridge, Mass.: Workers Compensation Research Institute, 2003.

Washington State Department of Labor and Industries, "Ortho and Neuro Surgeon Quality," undated Web page. As of July 13, 2007: http://www.lni.wa.gov/ClaimsIns/Providers/Research/OrthoNeuro

WCIRB—*see* Workers' Compensation Insurance Rating Bureau of California.

Weissert, William G., and Melissa Constable Musliner, “Case Mix Adjusted Nursing-Home Reimbursement: A Critical Review of the Evidence,” *The Milbank Quarterly*, Vol. 70, No. 3, 1992, pp. 455–490.

Wickizer Thomas M., Gary M. Franklin, Robert D. Mootz, Deborah Fulton-Kehoe, Roy Plaeger-Brockway, Diana Drylie, Judith A. Turner, and Terri Smith-Weller, “A Communitywide Intervention to Improve Outcomes and Reduce Disability Among Injured Workers in Washington State,” *The Milbank Quarterly*, Vol. 82, No. 3, September 2004, pp. 547–567.

Wickizer Thomas, Gary Franklin, Deborah Fulton-Kehoe, Jeremy Gluck, Robert Mootz, Terri Smith-Weller, and Rae Wu, *Centers of Occupational Health and Education: Final Report on Outcomes from the Initial Cohort of Injured Workers, 2003–2005*, Occupational Health Services Project, Washington State Department of Labor and Industries, April 22, 2007. As of August 1, 2007: http://www.lni.wa.gov/ClaimsIns/Files/Providers/ohs/CombinedReportApril2007.pdf

Workers’ Compensation Insurance Rating Bureau of California, “2002 California Workers’ Compensation Insured Losses and Expenses,” San Francisco, Calif.: WCIRB, June 2003.

———, “2003 California Workers’ Compensation Insured Losses and Expenses,” San Francisco, Calif.: WCIRB, June 28, 2004. As of August 7, 2007: https://wcirbonline.org/resources/data_reports/losses_expenses.html

———, “2004 California Workers’ Compensation Insured Losses and Expenses,” San Francisco, Calif.: WCIRB, June 2, 2005. As of August 7, 2007: https://wcirbonline.org/resources/data_reports/losses_expenses.html

———, “2005 California Workers’ Compensation Insured Losses and Expenses,” San Francisco, Calif.: WCIRB, June 23, 2006. As of August 7, 2007: https://wcirbonline.org/resources/data_reports/losses_expenses.html

———, “2006 California Workers’ Compensation Insured Losses and Expenses,” San Francisco, Calif.: WCIRB, June 18, 2007a. As of August 1, 2007: https://wcirbonline.org/resources/data_reports/losses_expenses.html

———, *WCIRB Summary of Policy Year Statistics—2007 Release*, San Francisco, Calif.: WCIRB, 2007b.

Wynn, Barbara O., Giacomo Bergamo, Rebecca Shaw, and Soeren Mattke, *Medical Care Provided California’s Injured Workers: An Overview of the Issues*, Santa Monica, Calif.: RAND Corporation, WR-394-ICJ, forthcoming.